

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:21:53 ; Search time 36 Seconds

(without alignments)
198,605 Million cell updates/sec

Title: US-09-944-403-42

Perfect score: 1325
Sequence: 1 MRPLVLLLLGLAAGSPPLD.....DSTFGFLVYSDMHSVPVFA 243

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1325	100.0	243	US-09-140-804-2	Sequence 2, Appl1
2	1325	100.0	243	US-09-336-536-3	Sequence 3, Appl1
3	1271	95.9	243	US-09-188-930-295	Sequence 295, Appl1
4	1258	94.9	243	US-09-336-536-10	Sequence 10, Appl1
5	1256	94.8	228	US-09-336-536-4	Sequence 4, Appl1
6	1200	90.6	228	US-09-336-536-11	Sequence 11, Appl1
7	675	50.9	128	US-09-336-536-7	Sequence 7, Appl1
8	650	49.1	128	US-09-336-536-14	Sequence 14, Appl1
9	433.5	32.7	247	US-08-463-911-2	Sequence 2, Appl1
10	428.5	32.3	247	US-09-140-804-8	Sequence 8, Appl1
11	428.5	32.3	247	US-09-118-408-3	Sequence 3, Appl1
12	428.5	32.3	247	US-09-506-855-3	Sequence 7, Appl1
13	424	32.0	244	US-08-463-911-7	Sequence 7, Appl1
14	424	32.0	244	US-09-140-804-3	Sequence 3, Appl1
15	424	32.0	244	US-09-336-536-20	Sequence 20, Appl1
16	424	32.0	244	US-09-336-536-1	Sequence 1, Appl1
17	417	31.5	231	US-09-530-423-2	Sequence 2, Appl1
18	412	31.1	246	US-08-463-911-4	Sequence 4, Appl1
19	399	30.1	294	US-09-188-930-294	Sequence 294, Appl1
20	347	26.2	60	US-09-336-536-6	Sequence 6, Appl1
21	324	24.5	60	US-09-336-536-13	Sequence 13, Appl1
22	316	23.8	433	US-08-383-744-2	Sequence 2, Appl1
23	316	23.8	433	US-08-989-336-2	Sequence 2, Appl1
24	316	23.8	433	PCT-US96-01427-2	Sequence 4, Appl1
25	280.5	21.2	215	US-09-140-804-4	Sequence 5, Appl1
26	280.5	21.2	215	US-09-140-804-5	Sequence 5, Appl1
27	277.5	20.9	185	US-08-463-911-3	Sequence 3, Appl1

28	276	20.8	236	US-09-140-804-6	Sequence 6, Appl1
29	266.5	20.1	198	US-09-188-930-138	Sequence 138, Appl1
30	257.5	19.4	130	US-09-485-316A-13	Sequence 13, Appl1
31	251.5	19.0	130	US-09-485-316A-12	Sequence 12, Appl1
32	243.5	18.4	130	US-09-485-316A-11	Sequence 11, Appl1
33	240	18.1	161	US-09-415-551-3	Sequence 3, Appl1
34	229.5	17.3	280	US-09-247-155-178	Sequence 178, Appl1
35	229.5	17.3	281	US-09-118-408-2	Sequence 2, Appl1
36	229.5	17.3	281	US-09-506-855-2	Sequence 2, Appl1
37	228	17.2	684	US-08-555-669-12	Sequence 12, Appl1
38	228	17.2	684	US-09-073-663-12	Sequence 12, Appl1
39	224	16.9	228	US-09-219-849-38	Sequence 38, Appl1
40	223	16.8	281	US-09-118-408-4	Sequence 44, Appl1
41	223	16.8	281	US-09-506-855-4	Sequence 44, Appl1
42	221	16.7	124	US-09-485-316A-9	Sequence 9, Appl1
43	219.5	16.6	1057	US-08-931-820-4	Sequence 4, Appl1
44	219.5	16.6	1078	US-08-963-825-21	Sequence 21, Appl1
45	219.5	16.6	1078	US-09-500-811-21	Sequence 21, Appl1

ALIGNMENTS

RESULT 1
US-09-140-804-2

Sequence 2, Application US/09140804

Patent No. 6197930

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS

FILE REFERENCE: 97-49

CURRENT APPLICATION NUMBER: US/09/140, 804

EARLIER FILING DATE: 1998-08-26

EARLIER APPLICATION NUMBER: 60/056, 983

NUMBER OF SEQ ID NOS: 47

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 243

TYPE: PRT

ORGANISM: Homo sapiens

US-09-140-804-2

Query Match

Best Local Similarity 100.0%; Score 1325; DB 4; Length 243;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGDGAPG 60

DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGDGAPG 60

QY 61 ARGEKGGRRPGLPEPRDPPGEGAGPAGTGAAGCSVPPRASFARKESSRRVPPSD 120

DB 61 ARGEKGGRRPGLPEPRDPPGEGAGPAGTGAAGCSVPPRASFARKESSRRVPPSD 120

QY 121 APLPFDRLVNEQGHYDAVTGKTCOVPGVYFAVHATVYRASLOPDLVNGKSTASFQ 180

DB 121 APLPFDRLVNEQGHYDAVTGKTCOVPGVYFAVHATVYRASLOPDLVNGKSTASFQ 180

QY 181 FFGGMPKRSASGAMRWLEPEDVWVQVGVDTIGTASIKTDSFTSGFLVYSDMHSRP 240

DB 181 FFGGMPKRSASGAMRWLEPEDVWVQVGVDTIGTASIKTDSFTSGFLVYSDMHSRP 240

QY 241 VFA 243

DB 241 VFA 243

RESULT 2

US-09-336-536-3

Sequence 3, Application US/09336536

Patent No. 6406884

NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentId Ver. 2.0
SEQ ID NO 4
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-4

Query Match 94.8%; Score 1256; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.8e-107;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SPPLDNDKIPSLCPGHPGLPGTPGHHGSOGLPRDGRDGRDGAAPGKGGKGRGRLPG 75
DB 1 SPPLDNDKIPSLCPGHPGLPGTPGHHGSOGLPRDGRDGRDGAAPGKGGKGRGRLPG 60
QY 76 PRDGPGRGAGAGPTGPAEGSVPRSAFSAKSESSESVPPSDAPLPFDRVLVNEQGH 135
DB 61 PRDGPGRGAGAGPTGPAEGSVPRSAFSAKSESSESVPPSDAPLPFDRVLVNEQGH 120
QY 136 YDAVTGKFTCOVGVYFAVHATVYRASLOFDLVNKGESIASFPQFGGMPKPSLSGGA 195
DB 121 YDAVTGKFTCOVGVYFAVHATVYRASLOFDLVNKGESIASFPQFGGMPKPSLSGGA 180
QY 196 MVRLEBEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSPVFA 243
DB 181 MVRLEBEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSPVFA 228

RESULT 6

US-09-336-536-11
Sequence 11, Application US/09336536

GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentId Ver. 2.0
SEQ ID NO 11
LENGTH: 228
TYPE: PRT
ORGANISM: Mus musculus
US-09-336-536-11

Query Match 90.6%; Score 1200; DB 4; Length 228;
Best Local Similarity 94.7%; Pred. No. 2.4e-102;
Matches 216; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 16 SPPLDNDKIPSLCPGHPGLPGTPGHHGSOGLPRDGRDGRDGAAPGKGGKGRGRLPG 75
DB 1 SPPLDNDKIPSLCPGHPGLPGTPGHHGSOGLPRDGRDGRDGAAPGKGGKGRGRLPG 60
QY 76 PRDGPGRGAGAGPTGPAEGSVPRSAFSAKSESSESVPPSDAPLPFDRVLVNEQGH 135
DB 61 PRDGPGRGAGAGPTGPAEGSVPRSAFSAKSESSESVPPSDAPLPFDRVLVNEQGH 120
QY 136 YDAVTGKFTCOVGVYFAVHATVYRASLOFDLVNKGESIASFPQFGGMPKPSLSGGA 195
DB 121 YDAVTGKFTCOVGVYFAVHATVYRASLOFDLVNKGESIASFPQFGGMPKPSLSGGA 180
QY 196 MVRLEBEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSPVFA 243
DB 181 MVRLEBEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSPVFA 228

RESULT 7

US-09-336-536-7
Sequence 7, Application US/09336536

Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentId Ver. 2.0
SEQ ID NO 7
LENGTH: 128
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-7

Query Match 50.9%; Score 675; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.2e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 AFSAKSESSESVPPSDAPLPFDRVLVNEQGHYDAVTGKFTCOVGVYFAVHATVYRASL 164
DB 1 AFSAKSESSESVPPSDAPLPFDRVLVNEQGHYDAVTGKFTCOVGVYFAVHATVYRASL 60
QY 165 QFDLVNKGESIASFPQFGGMPKPSLSGGA MVRLEBEDQVWVGVDYIGIYASIKTD 224
DB 61 QFDLVNKGESIASFPQFGGMPKPSLSGGA MVRLEBEDQVWVGVDYIGIYASIKTD 120
QY 225 STFSGFLV 232
DB 121 STFSGFLV 128

RESULT 8

US-09-336-536-14
Sequence 14, Application US/09336536

GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentId Ver. 2.0
SEQ ID NO 14
LENGTH: 128
TYPE: PRT
ORGANISM: Mus musculus
US-09-336-536-14

Query Match 49.1%; Score 650; DB 4; Length 128;
Best Local Similarity 94.5%; Pred. No. 2.3e-52;
Matches 121; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 105 AFSAKSESSESVPPSDAPLPFDRVLVNEQGHYDAVTGKFTCOVGVYFAVHATVYRASL 164
DB 1 AFSAKSESSESVPPSDAPLPFDRVLVNEQGHYDAVTGKFTCOVGVYFAVHATVYRASL 60
QY 165 QFDLVNKGESIASFPQFGGMPKPSLSGGA MVRLEBEDQVWVGVDYIGIYASIKTD 224
DB 61 QFDLVNKGESIASFPQFGGMPKPSLSGGA MVRLEBEDQVWVGVDYIGIYASIKTD 120
QY 225 STFSGFLV 232
DB 121 STFSGFLV 128

RESULT 9

US-08-463-911-2

```

; Sequence 2, Application US/08463911
; Patent No. 5869330
; GENERAL INFORMATION:
; APPLICANT: Scherer, Philipp E.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
; TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,911
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI95-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-911-2

```

```

Query Match      32.7%; Score 433.5; DB 2; Length 247;
Best Local Similarity 40.6%; Pred. No. 3,4e-32;
Matches 102; Conservative 32; Mismatches 92; Indels 25; Gaps 8;

QY 1 MRPLVLLLL-GLAAGSPPLDNNKIPSLCPGHPG-----LPCTPGHHSOGILPGRDGRG 54
DB 4 LQALLFLILPSHADDVTTTEELAPALVPPPKGTCAWMAGIPGHPGHNTPGRDGRD- 62
QY 55 RDGAPGAPGEGKGGPGLPGPRGD-----PGPRGAGPAGPTGPAGECSVPPRSA 105
DB 63 -----GTPGEGKEKGDAGILGPKGTGDVGMTGAGPRGFPOTGRKGPBEAAMTRSA 117
QY 106 FSAKRSERVPSPDAPLPEDRVLVNKGHYDAVTGKTCOVPGVYFAVAHATYRASLQ 165
DB 118 FSV-GLETRVTV-NVPIRFTKIFYNQNNHYDGSIGKFCYCNIPGLIYFSYHITVYMDVK 175
QY 166 FDLVNGESIASFPQFGGMPKPSASGAMVRLPEPDQVWVOY-GVGDYIGIYASIKTD 224
DB 176 VSLFKDKAVLFTYDQYQ-KNVDAQSGSVLHLLEVQDQWLVQYGDGDHNGLYADNVND 234
QY 225 STFGGLVYSD 235
DB 235 STFGGLVYSD 245

```

```

RESULT 10
US-09-140-804-8
; Sequence 8, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49

```

```

; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-140-804-8

```

```

Query Match      32.3%; Score 428.5; DB 4; Length 247;
Best Local Similarity 40.2%; Pred. No. 9,7e-32;
Matches 101; Conservative 32; Mismatches 93; Indels 25; Gaps 8;

QY 1 MRPLVLLLL-GLAAGSPPLDNNKIPSLCPGHPG-----LPCTPGHHSOGILPGRDGRG 54
DB 4 LQALLFLILPSHADDVTTTEELAPALVPPPKGTCAWMAGIPGHPGHNTPGRDGRD- 62
QY 55 RDGAPGAPGEGKGGPGLPGPRGD-----PGPRGAGPAGPTGPAGECSVPPRSA 105
DB 63 -----GTPGEGKEKGDAGILGPKGTGDVGMTGAGPRGFPOTGRKGPBEAAMTRSA 117
QY 106 FSAKRSERVPSPDAPLPEDRVLVNKGHYDAVTGKTCOVPGVYFAVAHATYRASLQ 165
DB 118 FSV-GLETRVTV-NVPIRFTKIFYNQNNHYDGSIGKFCYCNIPGLIYFSYHITVYMDVK 175
QY 166 FDLVNGESIASFPQFGGMPKPSASGAMVRLPEPDQVWVOY-GVGDYIGIYASIKTD 224
DB 176 VSLFKDKAVLFTYDQYQ-KNVDAQSGSVLHLLEVQDQWLVQYGDGDHNGLYADNVND 234
QY 225 STFGGLVYSD 235
DB 235 STFGGLVYSD 245

```

```

RESULT 11
US-09-118-408-3
; Sequence 3, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/118,408A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-118-408-3

```

```

Query Match      32.3%; Score 428.5; DB 4; Length 247;
Best Local Similarity 40.2%; Pred. No. 9,7e-32;
Matches 101; Conservative 32; Mismatches 93; Indels 25; Gaps 8;

QY 1 MRPLVLLLL-GLAAGSPPLDNNKIPSLCPGHPG-----LPCTPGHHSOGILPGRDGRG 54
DB 4 LQALLFLILPSHADDVTTTEELAPALVPPPKGTCAWMAGIPGHPGHNTPGRDGRD- 62
QY 55 RDGAPGAPGEGKGGPGLPGPRGD-----PGPRGAGPAGPTGPAGECSVPPRSA 105
DB 63 -----GTPGEGKEKGDAGILGPKGTGDVGMTGAGPRGFPOTGRKGPBEAAMTRSA 117
QY 106 FSAKRSERVPSPDAPLPEDRVLVNKGHYDAVTGKTCOVPGVYFAVAHATYRASLQ 165
DB 118 FSV-GLETRVTV-NVPIRFTKIFYNQNNHYDGSIGKFCYCNIPGLIYFSYHITVYMDVK 175

```



```
QY 6 VLLLLGLAAGSPPLDNNKIPSLCPGH-----PG-----LPCTPGHHGSQ 44
Db 7 VLLLLAL-----PCHDQETTTGGPGVLLPLPKGACTGMMAGIPGHPGN 50
QY 45 GLPGRDGDRDAGAPGEGKEGGRPGLPGPRGD-----PSPRGAGPAGPTGPA 95
Db 51 GAPGDRGD-----GTGEGKEGKDPGLIGPKDIDGTVGPAEGPRGFFGIQGRKEP 104
QY 96 GECVVPSPSAFSAKSESRRVPPSDAPLPFDRVLVNEQGHYDAVTGKFTCOVPGVYFAY 155
Db 105 GEGAVYRSASFV-GLFTYVITP-NMPIRTKIFYNQONHYDSTGKFGHCNIPGLYFAY 162
QY 156 HATVVRASLQFDLYKNGS-IASFQFGGWPKPASLSSGAMVRLPEPDQVWVQV-GVGD 213
Db 163 HITVVMQVKVSLFKDKAMLFYDQYQENNVDQA--SGSVLLHLFVGDQVWLQVYGEGB 220
QY 214 YIGIVASIKTDTSTGSGFLVYSD 235
Db 221 RNLGYADNDNDSTFTGFLYHYD 242
```

RESULT 15

```
US-09-336-536-20
; Sequence 20, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bosson, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 20
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-20
```

```
Query Match 32.0%; Score 424; DB 4; Length 244;
Best Local Similarity 40.5%; Pred. No. 2.5e-31;
Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;
```

```
QY 6 VLLLLGLAAGSPPLDNNKIPSLCPGH-----PG-----LPCTPGHHGSQ 44
Db 7 VLLLLAL-----PCHDQETTTGGPGVLLPLPKGACTGMMAGIPGHPGN 50
QY 45 GLPGRDGDRDAGAPGEGKEGGRPGLPGPRGD-----PSPRGAGPAGPTGPA 95
Db 51 GAPGDRGD-----GTGEGKEGKDPGLIGPKDIDGTVGPAEGPRGFFGIQGRKEP 104
QY 96 GECVVPSPSAFSAKSESRRVPPSDAPLPFDRVLVNEQGHYDAVTGKFTCOVPGVYFAY 155
Db 105 GEGAVYRSASFV-GLFTYVITP-NMPIRTKIFYNQONHYDSTGKFGHCNIPGLYFAY 162
QY 156 HATVVRASLQFDLYKNGS-IASFQFGGWPKPASLSSGAMVRLPEPDQVWVQV-GVGD 213
Db 163 HITVVMQVKVSLFKDKAMLFYDQYQENNVDQA--SGSVLLHLFVGDQVWLQVYGEGB 220
QY 214 YIGIVASIKTDTSTGSGFLVYSD 235
Db 221 RNLGYADNDNDSTFTGFLYHYD 242
```

Search completed: June 20, 2003, 11:26:49
Job time : 37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 06:14:37 ; Search time 3597 Seconds

(without alignments)
1141.107 Million cell updates/sec

Title: US-09-944-403-41

Perfect score: 1377
Sequence: 1 gactagctctcttgagctc.....aaaaaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb ba:*
2: gb htg:*
3: gb ln:*
4: gb om:*
5: gb ov:*
6: gb pat:*
7: gb ph:*
8: gb pl:*
9: gb pr:*
10: gb ro:*
11: gb sts:*
12: gb sy:*
13: gb un:*
14: gb vl:*
15: em ba:*
16: em fun:*
17: em hum:*
18: em ln:*
19: em mu:*
20: em om:*
21: em or:*
22: em ov:*
23: em pat:*
24: em ph:*
25: em pl:*
26: em ro:*
27: em sts:*
28: em un:*
29: em vl:*
30: em htg hum:*
31: em htg inv:*
32: em htg other:*
33: em htg mus:*
34: em htg pln:*
35: em htg rod:*
36: em htg mam:*
37: em htg yrt:*
38: em sy:*
39: em htgo hum:*
40: em htgo mus:*
41: em htgo other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1377	100.0	1377	6	AX464228
2	1347.4	97.9	1370	9	BC029485
3	1322.2	96.0	1337	9	AF329841
4	1322.2	96.0	1347	6	AR138193
5	1064.4	77.3	1068	9	HSM800923
6	909.8	66.1	191362	9	AP003396
7	909.8	66.1	219574	9	AP002956
8	896.2	65.1	182429	2	AP001557
9	696.2	50.6	4220	10	AF469650
10	694	50.4	173038	2	AP001003
11	693.6	50.4	1234	10	BC025174
12	691.2	50.2	1271	10	BC023068
13	612.8	44.5	198902	2	AP001156
14	513.4	37.3	173657	2	AC112557
15	504	36.6	729	6	AR138195
16	500.4	36.3	249283	2	AC124577
17	472.4	34.3	150653	2	AC107174
18	435.4	31.9	441	9	HUM2D92F04
19	187.4	13.6	150653	2	AC107174
20	161.8	11.8	2288	9	AK055132
21	107	7.8	1866	5	AB067770
22	100.6	7.3	4443	9	AK074129
23	100.6	7.3	4908	6	AX430953
24	100.6	7.3	128133	9	AL138787
25	100.6	7.3	173817	2	AC012141
26	99.6	7.2	855	6	AX039974
27	97.4	7.1	1938	9	HUM28CGE
28	92.6	6.7	152147	9	AC091842
29	92.6	6.7	153248	9	AC112191
30	91.6	6.7	1161	6	AX039965
31	91.6	6.7	1176	9	AF329836
32	91.6	6.7	2404	9	BC011699
33	91.4	6.6	1134	4	AF269230
34	90.6	6.6	947	10	BC028770
35	90.6	6.6	1276	6	AR034252
36	90.6	6.6	1276	6	AX195209
37	90.6	6.6	1276	6	AX358519
38	90.6	6.6	1276	6	AX358519
39	90.2	6.6	194674	2	AC122477
40	90	6.5	132884	2	AC128403
41	88.2	6.4	112554	2	AC129431
42	88	6.4	1357	6	AX068378
43	88	6.4	1393	9	AF329838
44	88	6.4	159766	9	AC104942
45	86.8	6.3	178850	9	AL359736

ALIGNMENTS

RESULT 1	AX464228	LOCUS	AX464228	DEFINITION	Sequence 361 from Patent WO0140466.	DNA	linear	PAT 16-JUL-2002
VERSION	AX464228	AX464228	AX464228	KEYWORDS	GI:21899130			
SOURCE	human.			ORGANISM	human.			
REFERENCE	1							
AUTHORS	Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, R., Gao, W.O., Gerritsen, M.E., Goddard, A., Godowski, P.J., Guirney, A.L., Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K.,							


```

/chromosome="11"
/map="11q23.3"
1.1337
/gene="CTRP5"
/note="ZS1G39"
198.929
/feature="CDS"

```

```

/product="complement-clq tumor necrosis factor-related
protein"
/protain_id="AAK17965.1"

```

BASE COUNT	235 a	417 c	459 g	226 t
ORIGIN				

Query Match	96.0%;	Score 1322.2;	DB 9;	Length 1337;
Best Local Similarity	99.8%;	Pred. No. 3.5e-218;		
Matches 1324;	Conservative	0;	Mismatches 3;	Indels 0;

QY	40	CGCGAGGAGCAACGAACCAAGGACTGGGGTGAACGGCAAGGACAGGGGGGGGACCTGGCCGGGAGGA	99
Db	11	CGAGAGGAGGAAACAGAGACTGGAGTGAACGGACAGGGGAGGGGGGGCTGGGCGGGAGGA	70
QY	100	AGGCGGGGGGCTGAGGACCAACCACTGAGAGGATCCGAGTGAAGCAGCGCCCGAAGAG	159
Db	71	AGCGCGGGGGCTGAGAGCAACCACTGAGAGGATCCGAGTGAAGCAGCGCCCGAAGAG	130
QY	160	GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGAGCTCCGAGCTCCGGTGC	219
Db	131	GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGAGCTCCGGTGC	190
QY	220	CAGGCTATGAGGCGCATCTCTGTCTGTCTGTCTCTGGGACTGTGGGGCGGCGGTGGCCCC	279
Db	191	CAGGCTATGAGGCGCATCTCTGTCTGTCTGTCTCTGGGACTGTGGGGCGGCGGTGGCCCC	250
QY	280	ACTGGAAGACAACAGATCCCGAGCTCTGCGCCGGGGGACCCCGGCTTCCAGACAGCC	339
Db	251	ACTGGAAGACAACAGATCCCGAGCTCTGCGCCGGGGGACCCCGGCTTCCAGACAGCC	310
QY	340	GAGCACCATGAGCAAGCTTGGCTTGGCCGGGCGCATGAGCGGACAGCCGACGGGCGC	399
Db	311	GAGCACCATGAGCAAGCTTGGCTTGGCCGGGCGCATGAGCGGACAGCCGACGGGCGC	370
QY	400	GCCGGGGGCTCCGGGAGAGAAAGCGGAGGGCGGAGGCTCGGACTTCCGGGACTTCGAAG	459
Db	371	GCCGGGGGCTCCGGGAGAGAAAGCGGAGGGCGGAGGCTCGGACTTCCGGGACTTCGAAG	430
QY	460	GGAACCCGGGGCGGAGGAGAGGCGGAGACCCGCGGGGGCCACCGGGCTGCGCGGGAGTG	519
Db	431	GGAACCCGGGGCGGAGGAGAGGCGGAGACCCGCGGGGGCCACCGGGCTGCGCGGGAGTG	490
QY	520	CTCGGTGCTCCGCGATCCGCTTCAAGCGCCAAAGCGCTCGAAGGCGGGTGCCTTCGCC	579
Db	491	CTCGGTGCTCCGCGATCCGCTTCAAGCGCGCCAAAGCGCTCGAAGGCGGGTGCCTTCGCC	550
QY	580	GTCTGACGCACTCTTGCTTTCGACCGCGGTGTGTGAACGAGCAGAGGACATTACGACGC	639
Db	551	GTCTGACGCACTCTTGCTTTCGACCGCGGTGTGTGAACGAGCAGAGGACATTACGACGC	610
QY	640	CGTCAACCGGCAAGTCACTGCGAGGTGCTGGGGCTTACTACTTGGCGTCATGCAC	699
Db	611	CGTCAACCGGCAAGTCACTGCGAGGTGCTGGGGCTTACTACTTGGCGTCATGCAC	670
QY	700	CGTCAACCGGCGAGCTGCAAGTTTATCTGTGAAGATGCGAAATCCATTGCTCTTT	759
Db	671	CGTCAACCGGCGAGCTGCAAGTTTATCTGTGAAGATGCGAAATCCATTGCTCTTT	730
QY	760	CTTCAAGTTTTCGGGGGAGGCCCAAGCCAGCTGCTTCGGGGGGGGCAATGATGAG	819

Db	731	CTTCAGATTTTTCGGGGGGGTGGCCCAAGCCAGCTCCCTCTCGGGGGGGGCGCATGTGTGAG	790
Qy	820	GCTGGAGCCCTGAGGAGCCAAAGTGTGGGTGCAGGTGGGTGTGTGGTGACTACATTTGGCATCTTA	879
Db	791	GCTGGAGCCCTGAGGAGCCAAAGTGTGGGTGCAGGTGGGTGTGTGGTGACTACATTTGGCATCTTA	850
Qy	880	TGGCAGGATCAAGACAGACAGACACTTCTCCGGATTTCTGGTGTACTCCGACTGGACAG	939
Db	851	TGCCAGCATACAGACAGACAGACACTTCTCCGGATTTCTGGTGTACTCCGACTGGACAG	910
Qy	940	CTCCCAAGTCTTGTCTTAGTGTGCCCATCTGCATAAGTGAAGTCACTGCTCACTCTTAAGAGG	999
Db	911	CTCCCAAGTCTTGTCTTAGTGTGCCCATCTGCATAAGTGAAGTCACTGCTCACTCTTAAGAGG	970
Qy	1000	AGGGTGAAGGCTGACAAACAGGTCATCAGAGAGGCTGGCCCCCTCGAATATTTGTGAA	1058
Db	971	AGGGTGAAGGCTGACAAACAGGTCATCAGAGAGGCTGGCCCCCTCGAATATTTGTGAA	1033
Qy	1060	TGACTAAGGAGGTGGGGGTAGAGCACTCTCCGTCTGTGCTGGCAGAGGAATGGGAACGT	1115
Db	1031	TGACTAAGGAGGTGGGGGTAGAGCACTCTCCGTCTGTGCTGGCAGAGGAATGGGAACGT	1090
Qy	1120	GGCTGTCTGCACATCAAGTCTGGCAGCAAGGGGAGTGGCTGGAATTTCTGCCCAAGACAG	1172
Db	1091	GGCTGTCTGCACATCAAGTCTGGCAGCAAGGGGAGTGGCTGGAATTTCTGCCCAAGACAG	1155
Qy	1180	AGGATGTGCTGTGTCTGGCAAGTGTAAAGTCCCGCAGTTGTCTGTGTCAGAGGCCACAGG	1233
Db	1151	AGGATGTGCTGTGTGTCTGGCAAGTGTAAAGTCCCGCAGTTGTCTGTGTCAGAGGCCACAGG	1210
Qy	1240	TGGGATGCTCTTCTCTGTGCTCTGTGCTCTCTGTGAATCTCCCGACCCCTCTGTCTCTCT	1299
Db	1211	TGGGATGCTCTTCTCTGTGCTCTGTGCTCTCTGTGAATCTCCCGACCCCTCTGTCTCTCT	1272
Qy	1300	GGGGCGGGCCCTTTTCTCAGAGATCACTCAATTAACCTTAAGAACCTTCATATAAAAAAAA	1355
Db	1271	GGGGCGGGCCCTTTTCTCAGAGATCACTCAATTAACCTTAAGAACCTTCATATAAAAAAAA	1330
Qy	1360	AAAAAAAA 1366	
Db	1331	AAAAAAAA 1337	

RESULT 4			
ARI38193			
LOCUS	ARI38193	1347 bp	DNA
DEFINITION	Sequence 1 from patent US 6197930.		linear
ACCESSION	ARI38193		
VERSION	ARI38193.1	GI:14479702	
KEYWORDS			
SOURCE	Unknown.		

REFERENCE	1 (bases 1 to 1347)
AUTHORS	Sheppard, P. O. and Humes, J. M.
TITLE	Adipocyte-specific protein homologs
JOURNAL	Patent: US 6197930-A 1 06-MAR-2001;
FEATURES	location/Qualifiers
source	1..1347

```

/organism="unknown"
BASE COUNT      235 a      421 c      465 g
ORIGIN

```

Query Match	96.04	Score 1322.2	DB 6	length 1347
Best Local Similarity	99.88	Pred. No. 3.5e-218		
Matches 1324, Conservative	0	Mismatches 3	Indels 0	Gaps 0

Oy 40 CGG CAG GAG GAG CGA ACG AAG CTG GGG GTG ACG GCA GGG GCG CTC GCG GGG AGA 99
|||
Db 11 CGA GAG GAG CGA ACG AAG CTG GGG GTG ACG GCA GGG GCG CTC GCG GGG AGA 70
Oy 100 AGG CCG GGG CTT GAG AGC ACC ACG ACG TCG GAG GTG ACG ACG CCG CCG AAG AGG 159

QY 299 CCCAGGCTCTGCCCCGGGGGCAACCCGGGCTTCCAGGCAAGCCGGGCAACGAGCCAG 358
 Db 1 CCCAGGCTCTGCCCCGGGGGCAACCCGGGCTTCCAGGCAAGCCGGGCAACGAGCCAG 60
 QY 359 GGGCTTGGCCGGGCGCGCATGCGCGGCAAGCGCGGCAAGCGCGGCGCGGCGCGGCGCGG 418
 Db 61 GGGCTTGGCCGGGCGCGCATGCGCGGCAAGCGCGGCGCGGCGCGGCGCGGCGCGGCGG 120
 QY 419 AAAAGGAGAGGCGCGGAGGCGCGGAGCTGCCGGGAGCTCCAGAGGAGACCCGGGCGCGAGGA 478
 Db 121 AAAAGGAGAGGCGCGGAGGCGCGGAGCTGCCGGGAGCTCCAGAGGAGACCCGGGCGCGAGGA 180
 QY 479 GAGGGGGGAGCGCGCGGGGCGCGGAGCTGCCGGGAGCTCCGGTGGCTCCCGGATCC 538
 Db 181 GAGGGGGGAGCGCGCGGGGCGCGGAGCTGCCGGGAGCTCCGGTGGCTCCCGGATCC 240
 QY 539 GCGTTTCAGGCGCAAGCGCTCCGAGAGCGGGGTGGCTCCGCGGCTGAGAGCACTTTCGCC 598
 Db 241 GCGTTTCAGGCGCAAGCGCTCCGAGAGCGGGGTGGCTCCGCGGCTGAGAGCACTTTCGCC 300
 QY 599 TTGCAACCGGCTGCTGTGTAAGCAAGAGGAGCAATTAAGACCGGTCACCGGCAAGTTCAAC 658
 Db 301 TTGCAACCGGCTGCTGTGTAAGCAAGAGGAGCAATTAAGACCGGTCACCGGCAAGTTCAAC 360
 QY 659 TGCCAGGTCCTGGGGGCTACTACTTCCGCGTCAGACCGGTCACCGGTCACCGGTCACCGGTC 718
 Db 361 TGCCAGGTCCTGGGGGCTACTACTTCCGCGTCAGACCGGTCACCGGTCACCGGTCACCGGTC 420
 QY 719 CAGTTTGAATCTGTGTAAGAGATGCGCAATTCATTCGCTCTTCAGTTTTCGGGGGG 778
 Db 421 CAGTTTGAATCTGTGTAAGAGATGCGCAATTCATTCGCTCTTCAGTTTTCGGGGGG 480
 QY 779 TGCCCAAGGCGGCTGCTCTCCGGGGGGGCGAGGTCGAGGCTGAGACCTGAGAGCA 838
 Db 481 TGCCCAAGGCGGCTGCTCTCCGGGGGGGCGAGGTCGAGGCTGAGACCTGAGAGCA 540
 QY 839 GTGTGGGTGCAAGGTGGGTGTGGGTGTAATTCATTCGAGGTCGAGGTCGAGGTCGAGGTC 898
 Db 541 GTGTGGGTGCAAGGTGGGTGTGGGTGTAATTCATTCGAGGTCGAGGTCGAGGTCGAGGTC 600
 QY 899 AGCACTTCTCTCGGATTTCTGTGTACTCCGACAGCTCCGACAGTCTTTCCTTAC 958
 Db 601 AGCACTTCTCTCGGATTTCTGTGTACTCCGACAGCTCCGACAGTCTTTCCTTAC 660
 QY 959 TGCCCACTGCAAGAGGAGCTGACTCTCACTCCAGAGAGGTCGAGGTCGAGGTCGAGGTC 1018
 Db 661 TGCCCACTGCAAGAGGAGCTGACTCTCACTCCAGAGAGGTCGAGGTCGAGGTCGAGGTC 720
 QY 1019 CAGGTGATCCAGAGAGGCTGGCCCCCTGGAATATTGTGAATGACTAGAGGAGTGGGGTA 1078
 Db 721 CAGGTGATCCAGAGAGGCTGGCCCCCTGGAATATTGTGAATGACTAGAGGAGTGGGGTA 780
 QY 1079 GAGCACTCTCGTCTGCTGTGTGTGCAAGAGATGAGGATGAGGTCGAGTGTGAGTCAAGGTC 1138
 Db 781 GAGCACTCTCGTCTGCTGTGTGTGCAAGAGATGAGGATGAGGTCGAGTGTGAGTCAAGGTC 840
 QY 1139 TGCGAGCATGAGGAGTGGGCTGGAATTTCTGCCCAAGACAGAGAGAGTGTGCTGAGGTC 1198
 Db 841 TGCGAGCATGAGGAGTGGGCTGGAATTTCTGCCCAAGACAGAGAGAGTGTGCTGAGGTC 900
 QY 1199 AAGGTATAGTCCCGCAAGTTGCTGTGTGTGCAAGAGCCGAGGTGGGGTCTCTTCTCTGG 1258
 Db 901 AAGGTATAGTCCCGCAAGTTGCTGTGTGTGCAAGAGCCGAGGTGGGGTCTCTTCTCTGG 960
 QY 1259 TCTCTGCTCTCTGAGATCTCCCAAGGCTCTGAGGTCGAGGTCGAGGTCGAGGTCGAGGTC 1318
 Db 961 TCTCTGCTCTCTGAGATCTCCCAAGGCTCTGAGGTCGAGGTCGAGGTCGAGGTCGAGGTC 1020
 QY 1319 GAGATCACTCAATTAAGCAAGCACTGATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1364
 Db 1021 GAGATCACTCAATTAAGCAAGCACTGATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1066

RESULT 6
 AP003396/c
 LOCUS
 DEFINITION
 Homo sapiens genomic DNA, chromosome 11q clone:RP11-334B6, complete
 sequences.
 ACCESSION
 AP003396
 VERSION
 AP003396.1
 KEYWORDS
 GI:20302607
 SOURCE
 HTG.
 ORGANISM
 Homo sapiens
 Homo sapiens DNM, clone:RP11-334B6.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Torok,Y., Matanabe,H. and Sakaki,Y.
 Human sapiens genomic DNA
 Published Only in Database (2002)
 2 (bases 1 to 191362)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Torok,Y., Matanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (12-MAR-2001) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel:81-45-503-9111, Fax:81-45-503-9170
 This work was done in collaboration with Arai, Y., Kubo, T. and
 Ohki, M.
 National Cancer Center Research Institute
 Cancer Genomic Division
 5-1-1, Tsukiji, Chuo-ku, Tokyo, JAPAN
 zip: 104-0045
 phone: 81-3-3542-2511 ex 4752, fax: 81-3-3542-0688 e-mail:
 yara@nc.90.jp.
 yara@nc.90.jp.
 FEATURES
 source
 1.191362
 /location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="RP11-334B6"
 BASE COUNT
 46535 a 45977 c 48722 g 50128 t
 ORIGIN
 Query Match 66.1%; Score 909.8; DB 9; Length 191362;
 Best Local Similarity 99.8%; Pred. No. 2,6e-147;
 Matches 911; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 438 CGGAGCTGCCGGGAGGAGCTCCAGAGGAGGAGCCCGGGCGCGAGAGAGCGCGGAGC 497
 Db 6384 GAGGAGCTGCCGGGAGGAGCTCCAGAGGAGGAGCCCGGGCGCGAGAGAGCGCGGAGC 6325
 QY 498 CCAACGGGCTGCGCGGAGAGTGTGCGGATCCGAGATCCGCTTCAAGCGCAAGCGCT 557
 Db 6324 CCAACGGGCTGCGCGGAGAGTGTGCGGATCCGAGATCCGCTTCAAGCGCAAGCGCT 6265
 QY 558 CCGAGAGCGGGGCTGCGCGGAGAGTGTGCGGATCCGAGATCCGCTTCAAGCGCGTGTGA 617
 Db 6264 CCGAGAGCGGGGCTGCGCGGAGAGTGTGCGGATCCGAGATCCGCTTCAAGCGCGTGTGA 6205
 QY 618 ACCAGAGGAGCAATTAAGAGCGCGTCAAGCGGCAAGTTCACTGCGCAGGTCGAGGTC 677
 Db 6204 ACCAGAGGAGCAATTAAGAGCGCGTCAAGCGGCAAGTTCACTGCGCAGGTCGAGGTC 6145
 QY 678 ACTACTTCCGCGCTGCAATGCAAGCGCTTACCGGGCCAGCGCTGCAATTTGATCTGTGA 737
 Db 6144 ACTACTTCCGCGCTGCAATGCAAGCGCTTACCGGGCCAGCGCTGCAATTTGATCTGTGA 6085
 QY 738 ATGGCAATTCATTTGCTCTTTCAGTTTTCGAGGAGGTCGAGGTCGAGGTCGAGGTC 797
 Db 6084 ATGGCAATTCATTTGCTCTTTCAGTTTTCGAGGAGGTCGAGGTCGAGGTCGAGGTC 6025
 QY 798 TCTCGGGGGGGGCGATGTGAGGCTGAGAGCCAGAGTGTGAGTGTGAGTGTGAGTGTG 857

Db	Accession	Version	Keywords	Source	Organism	Reference Authors	Title Journal Authors	Comment
Db	6024	TC006666	GGGCGCATGTGAGAGGCTGAGACCTGAGAGCCAACTGTGGGTGAGGTGGT					
Qy	858	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5964	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	918	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5904	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	918	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5844	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	1038	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5784	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	1038	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5724	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	1158	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5664	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	1218	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5604	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	1278	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5544	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Qy	1338	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
Db	5484	TGGGTGACATCAATTTGGCATCTATATGCCAGATCAAGAAGACGACCTTTCTCCGATTTTC						
RESULT 7	AP002956/c							
LOCUS	AP002956	219574 bp	DNA	linear	PRI 15-AUG-2002			
DEFINITION	Homo sapiens genomic DNA, chromosome 11q clone:RI105H09, complete							
ACCESSION	AP002956							
VERSION	AP002956.1	GI:22255355						
KEYWORDS	HTG.							
SOURCE	Homo sapiens DNA, clone:RI105H09.							
ORGANISM	Homo sapiens							
REFERENCE	Mumukshu, Eutheria; Primates; Catarrhini; Hominiidae; Homo.							
AUTHORS	Hattori, M., Toyoda, A., Taylor, T.D., Fujiyama, A., Yada, T.,							
TITLE	Homo sapiens 219,574 genomic DNA of 11q							
JOURNAL	Published Only in Database (2002)							
AUTHORS	Hattori, M., Toyoda, A., Taylor, T.D., Fujiyama, A., Yada, T.,							
TITLE	Direct Submission							
JOURNAL	Submitted (24-NOV-2000) Masahira Hattori, The Institute of Physical							
COMMENT	and Chemical Research (RIKEN), Genomic Sciences Center (GSC),							
	1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan (GSC),							
	(E-mail: hattori@gs.c.riken.go.jp, URL: http://hgp.gs.c.riken.go.jp/,							
	Te: 81-45-503-9111, Fax: 81-45-503-9170)							
	Overlapping Clones: pD11A4, pQ3868							
	Cytogenetic Position: 21q22.1, region: D1S226-AML							
	STS Markers (ePCR): stsG53747, SHG-16045, D1S2450, SHG-7143,							
	WT-16056							
	stsG31054, sts-X57110, WT-8978, GDB:198117, stsG39284, SHG-130657,							
	D1S1941E, WT-14589, stsG60191, stsG50535, A002048, stsG26946,							
	stsG41092,							

TIGR-A005D38,	Bdaa7902,	D1186869E,	RH80030
Contamination: none detected			
non-AGCT bases: none			
Additional author information			
Arai,Y., Kubo,T.,Ohki,M.			
National Cancer Center Research Institute			
Cancer Genomic Division			
5-1-1, Tsukiji1, Chuo-Ku, Tokyo, JAPAN			
zip: 104-0045			
phone: +81-3-3542-2511 ex4752, fax: +81-3-3542-0688 e-mail:			
yara@ncc.go.jp.			
Location/Qualifiers			
1..219574			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/chromosome="11"			
/map="11q"			
/clone="R1105h09"			
BASE COUNT	54883 a	50356 c	53844 g 60491 t
ORIGIN			
Query Match	Best Local Similarity	96.1%; Score 909.8; DB 9; Length 219574;	
Matches 911; Conservative	0; Mismatched	2; Indels	0; Gaps 0;
QY	438 CGGAGCTGC	CGGGACCTCGAGGGGACC	CGCGCCGGAGAGAGCGCGGACCCCGGGGCC 497
DB	209122 CAGAGCTGC	CGGACCTCGAGGGGACC	CGCGCGCGGAGAGAGCGCGGACCCCGGGGCC 20906
QY	498 CCACCGGCGCTCG	CGGGAGTGTCTCGTGCTTC	CGGATCCGGCTTCAGGCGCAAGGCT 557
DB	209062 CCACCGGCGCTCG	CGGGAGTGTCTCGTGCTTC	CGGATCCGGCTTCAGGCGCAAGGCT 20900
QY	558 CCGAAGCGCGGTGTCT	CGCGCGTCTGAAGCACCTT	GCCCTTGACCCGCGTCTGTGA 617
DB	209002 CCGAAGCGCGGTGTCT	CGCGCGTCTGAAGCACCTT	GCCCTTGACCCGCGTCTGTGA 20894
QY	618 ACGAGCAGGAGCATTA	ACGAGCGCGGTCAACCGGCA	AGTTCACTCCGCGAGTCCCTGGGGCT 677
DB	208942 ACGAGCAGGAGCATTA	ACGAGCGCGGTCAACCGGCA	AGTTCACTCCGCGAGTCCCTGGGGCT 20888
QY	678 ACTACTTGCGCGGTCA	TGTCATCGGCGCACCTG	CACGATTGATCTGTGTAAGA 737
DB	208882 ACTACTTGCGCGGTCA	TGTCATCGGCGCACCTG	CACGATTGATCTGTGTAAGA 20882
QY	738 ATGGGAAATCCAATTC	CGCTCTTCTTCCAGTTTTT	TGGGGGGGTGGCCCCAAGCCAGCTTCG 757
DB	208822 ATGGGAAATCCAATTC	CGCTCTTCTTCCAGTTTTT	TGGGGGGGTGGCCCCAAGCCAGCTTCG 20876
QY	758 TCTCGGGGGGGGCGCAT	TGTGTGAGGCTGAGGCTG	AGAGCCAAAGTGTGGGTGGGTG 857
DB	208762 TCTCGGGGGGGGCGCAT	TGTGTGAGGCTGAGGCTG	AGAGCCAAAGTGTGGGTGGGTG 20870
QY	858 TGGGTGATCTAATTGG	GCATCTATGCGCAGATCA	AAGACAGACAGACCTTCTCCGAAATTC 917
DB	208702 TGGGTGATCTAATTGG	GCATCTATGCGCAGATCA	AAGACAGACAGACCTTCTCCGAAATTC 20864
QY	918 TGATGTACTCCGATCT	GCGACAGCTCCCAGATCTT	TGCTTAGAGCCCACTGCAAAAGGAG 977
DB	208642 TGATGTACTCCGATCT	GCGACAGCTCCCAGATCTT	TGCTTAGAGCCCACTGCAAAAGGAG 20858
QY	978 TCATGCTCTCACTCT	TAGAAGAGGGGTGAGG	CTGACACAACCAAGATCATCCAGAGGGCT 1037
DB	208582 TCATGCTCTCACTCT	TAGAAGAGGGGTGAGG	CTGACACAACCAAGATCATCCAGAGGGCT 20852
QY	1038 GGCCCCCTGGAATTTG	TGGAATGCTAGGAGAGGT	GGGTATAGACATCTCCGTCCGTGCT 1097
DB	208522 GGCCCCCTGGAATTTG	TGGAATGCTAGGAGAGGT	GGGTATAGACATCTCCGTCCGTGCT 20846
QY	1098 GCTGGCAAGAAATGG	GAACAGTGGCTGTCTG	GATCAGGTCTGGAGCAATGGGGCAGTGG 1157
DB	208462 GCTGGCAAGAAATGG	GAACAGTGGCTGTCTG	GATCAGGTCTGGAGCAATGGGGCAGTGG 20840

```

QY 1158 CTGATTTCTGCGCAAGAGAGTGTCTGTGTCGAGCAAGTAAAGTCCCAAGTT 1217
DB 208402 CTGATTTCTGCGCAAGAGAGTGTCTGTGTCGAGCAAGTAAAGTCCCAAGTT 208343
QY 1218 GCTGTGTGTCAGAGAGCCAGCGGTGGGTGTCTCTTCTGTCCTCTGCTCTCTGATC 1277
DB 208342 GCTGTGTGTCAGAGAGCCAGCGGTGGGTGTCTCTTCTGTCCTCTGCTCTCTGATC 208283
QY 1278 CTCCCAACCCCTCTGCTGTCGAGGCGCGCCCTTTTCTCAGAGATCACTCAATTAACCT 1337
DB 208282 CTCCCAACCCCTCTGCTGTCGAGGCGCGCCCTTTTCTCAGAGATCACTCAATTAACCT 208223
QY 1338 AAGAACCTCATTA 1350
DB 208222 AAGAACCTCTCTTA 208210

RESULT 8
AP001557 182429 bp DNA linear HTG 03-NOV-2000
LOCUS Homo sapiens chromosome 11 clone RP11-680A7 map 11q23, WORKING
DEFINITION DRAFT SEQUENCE, 21 unordered pieces.
ACCESSION AP001557
VERSION AP001557.3 GI:11094164
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-680A7.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 182429)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Toroki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens 182,429 genomic DNA of 11q23
JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 182429)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Toroki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suenho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel:81-45-503-9111, Fax:81-45-503-9170)
URL:http://hgp.gsc.riken.go.jp/,
On Nov 3, 2000 this sequence version replaced gi:8117391.
COMMENT ----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: Humdrat11
Center clone name: RP11-680A7
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 175597 bases at least Q40
Consensus quality: 178555 bases at least Q30
Consensus quality: 179768 bases at least Q20
Insert size: 180429; sum-of-coverage
Quality coverage: 9.01x in Q20 bases; sum-of-coverage
NOTE: This is a 'working draft' sequence. It currently consists of
21 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1 30107 contig of 30107 bp in length
30208 54357 contig of 24150 bp in length
54458 69839 contig of 15382 bp in length
69940 84048 contig of 14109 bp in length
84149 95177 contig of 11025 bp in length
95178 95277 contig of 100 bp in length
95278 107763 contig of 1486 bp in length
107764 107863 contig of 100 bp
107864 119057 contig of 11194 bp in length
119058 119157 contig of 100 bp
119158 128134 contig of 8977 bp in length
128135 128234 contig of 100 bp
128235 136903 contig of 8669 bp in length
136904 137003 contig of 100 bp
137004 145887 contig of 8884 bp in length
145888 145987 contig of 100 bp
145988 153199 contig of 7212 bp in length
153200 153299 contig of 100 bp
153300 159565 contig of 6266 bp in length
159566 159665 contig of 100 bp
159666 163580 contig of 3915 bp in length
163581 163680 contig of 100 bp
163681 165152 contig of 1472 bp in length
165153 165252 contig of 100 bp
165253 166901 contig of 1649 bp in length
166902 167001 contig of 100 bp
167002 170460 contig of 3459 bp in length
170461 170560 contig of 100 bp
170561 172761 contig of 2201 bp in length
172762 172861 contig of 100 bp
172862 175605 contig of 2744 bp in length
175606 175705 contig of 100 bp
175706 178308 contig of 2603 bp in length
178309 178408 contig of 100 bp
178409 180918 contig of 2510 bp in length
180919 181018 contig of 100 bp
181019 182429 contig of 1411 bp in length.
Location/Qualifiers.
1. 182429
/organism="Homo sapiens"
/db xref="taxon:9606"
/chromosome="11"
/map="11q23"
/clone="RP11-680A7"

```

FEATURES

source

	Query Match Similarity	95.1%	Score 896.2	DB 2	length 182429
	Best Local Similarity	99.6%	Pred. No.5.8e-145		
	Matches 909	Conservative 0	Mismatches 13	Indels 1	Gaps 1
QY	438	CGGAGCTGCGGGAGACCTTCAGAGGAGACCCCGGCGCGGAGAGAGAGCGGAGACCCCGCGGGGC	497		
Db	108011	CAGAGCTGCGGGAGACTTCGAGGGAGACCCCGGCGCGGAGAGAGAGCGGAGACCCCGCGGGGC	108070		
QY	498	CCACCCGGGCTCTGCCGGGGAGTGTCTGGTGTGCTTCGCGATCCGCTTTCAGCGCCAGCGCT	557		
Db	108071	CCACCCGGGCTCTGCCGGGGAGTGTCTGGTGTGCTTCGCGATCCGCTTTCAGCGCCAGCGCT	108130		
QY	558	CCGAGAGCCGGGTGCTCTCGCGCGTCTGACGCAACCCCTTGCCCTTGACCCCGGTGCTGTGA	617		
Db	108131	CCGAGAGCCGGGTGCTCTCGCGCGTCTGACGCAACCCCTTGCCCTTGACCCCGGTGCTGTGA	108190		
QY	618	ACGACGAGGAGCAATTACGACGCGCGTCAACCGGAGAGTTACCTGCGACAGGTGCTCGAGTCT	677		
Db	108191	ACGAGCA-GGACATTACGACGCGCGTCAACCGGAGAGTTACCTGCGACAGGTGCTCGAGTCT	108249		
QY	678	ACTACTTGGCGGTCCATGACGACCGGTCTACCGGGCCAGCCTGCAATTGATCTGTGTGAAGA	737		
Db	108250	ACTACTTGGCGGTCCATGACGACCGGTCTACCGGGCCAGCCTGCAATTGATCTGTGTGAAGA	108309		
QY	738	ATGGCGAATCAATTCCTCTTCTTCCAGTTTCTTGGGGGGGGGGGCCAAGCCAGCCTCGC	797		
Db	108310	ATGGCGAATCAATTCCTCTTCTTCCAGTTTCTTGGGGGGGGGGGCCAAGCCAGCCTCGC	108369		

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="9"
/map="26 CM"

```

```

gene
  52. 1806
  /gene="Mfrp"
  CDS
    52. 1806
    /gene="Mfrp"
    /note="MFRP"
    /codon_start=1
    /product="membrane-type frizzled-related protein"
    /protein_id="AA089216.1"
    /db_xref="GI:22023767"
    /translation="MKDYDVILPEASELSKTEFCNPAFPEAGSPCPALQDVG
    SRLQAPMAORLRLGDPDCHSMFCILLGILLGLAVIAQLOASLPRTK
    NPILTRGLTPMGVLPSTPTPTTTTTPARQOQDAASPTHQTCGLLPEPSF
    PSEPNPDLPLPSHCWHIQAAGOTIQLKQALSTESMLTCLPDLBITSETGEL
    LRVGKTPPATLANTSHLRVFSVSDNDVBSGQAWYQAVAPHMSCANEFCDIL
    LCLKRDVDCGITECADGSDANCSAKTLGGLVGFSTPNYPHYPOHQLCT
    WYLEVPVYGRILEFNFSLQAQAEKDYVEVYEAENLGTSEFLGFCAPPLANV
    SSMHQLAVIFKTDLGISGGFLATYQALNTTESGCPAAEFCOSGYRDLQMDLMD
    CANDSNDCSHLSRQPLDCEPQVEMCLSYNTTAPFPIWYGLATOTREVDILNG
    YKSLTSLPCTQTFORFLGILLVPRCTSLGTLTPPCNSVCAAEQCCSIALACTPMP
    FNCRLPVAASLEACSQP"
  3097. 3828
  /gene="Clqtnf5"
  3097. 3828
  /gene="Clqtnf5"
  /note="Clqtnf5"
  /codon_start=1
  /product="complement-clq tumor necrosis factor-related
  protein"
  /protein_id="AA089217.1"
  /db_xref="GI:22023768"
  /translation="WRPLALALLGLVSGPPLDNDKIPLSLPEQPLPPTPGHNSQ
  GLPDRDARDGAPGAGEKGGRRPGLPDRPBPRLGAGMGWAIQPAEGCSVP
  RSASARSRSESRVPPADTLPDRVLINEGHYPDTGFTCCVPEVYFAVATYV
  RASLOPLVKNKOSIASFPQYFGMPKPASISGGAMVRLBEPDQWVQVGVYIGLY
  ASIKTDSFESFLVSDMHSPPVA"
BASE COUNT      873 a      1282 c      1147 g      918 t
ORIGIN
Query Match      50.6%; Score 696.2; DB 10; Length 4220;
Best Local Similarity 78.0%; Pred. No. 2.3e-110;
Matches 921; Conservative 0; Mismatches 238; Indels 22; Gaps 6;
QY 189 GAGAGAACCCCGGCGTCCGCGCTCCGCGTCCAGCGCTATGAGGCCACTCTCTGCTGC 248
DB 3059 GAGAGATCTCTGAGTCTGAGCTCCGCGGCTCACCAACATGAGGCACTTCTGCTCTC 3118
QY 249 TGCTCTGGGCGTGGCGCGCGCGCTCGCCCCCACTGAGACAGACAAGATCCCAAGCCCT 308
DB 3119 TGCTCTGGGCGTGGTGTCAAGCTCTCTCTCTGAGACAGACAAGATCCCAAGCCCTGT 3178
QY 309 GCCCGGGGACCCCGGCTTCCAGGCAAGCGCGGCAACATGAGGAGCGGCTTGC 368
DB 3179 GTCCCGGGGACCCCGGCTTCCAGGCAACCAAGGTCAACATGAGGAGCGGCTTGC 3238
QY 369 GCCCGGATGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 428
DB 3239 GCCGTGACGGCGCTGATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3298
QY 429 GCGGAGGCGCGGACTGCGCGGAGACTGAGGGGAGCCCGGCGCGGAGAGAGCGGAG 488
DB 3299 GCGGAGAGCCGGGACTTACCTGCGCGCACTGAGGAGCGCGGCGCGGAGAGAGCGG 3358
QY 489 CCGCGGGGCGCAACCGGGGCTGCGCGGAGTGTCTGCGTCTCCGCGATCCGCTTCA 548
DB 3359 CCATGGGGGCTATCGGGGCTGCGGGGAGTGTCTGCGTACCCCAAGATCAAGCTTCA 3418
QY 549 CCAAGCGCTCCGAGAGCGCGGGTGTCTCGCGGCTGAGCGACCTTGGCGCTTCA 608
DB 3419 CCAAGCGATCCGAGAGCGCGGGTACTTCCGCGAGCGGAGACACCCCTTCACTTCA 3478
QY 609 TGCTGTGAGAGAGGAGGACATTAAGAGCGCGGTACCGGAGGAGTTCACTGCGAGTGC 668
DB 3479 TGCTGTAAATAGAGAGGCGCTTACGACCCCACTACTGCGAGTTCACTGCGAGTGC 3538

```

```

QY 669 CTGGGCTTACTACTTCCCGCTCCATGCCACCGCTTACCGGGCGAGCTGCACTTGCATC 728
DB 3539 CTGGCGCTACTACTTCTTGTCTGTCTGACGCCACTCTCTACCGGGCGAGCTTGCATTTGATC 3598
QY 729 TGGTGAAGATGCGGATATCATGCGCTTTCTTCCAGTTTTCGCGGGGCTGCGCCCAAGC 788
DB 3599 TTGTCAAAAAGCGGAGCTCATGCGCTTCTTCCAGTATTTTGGGGGCTGCGCCCAAGC 3658
QY 789 CAGCTGCTCTCGGGGCGGGCCATGAGTGGAGCTGAGCGCTGAGACCAAGTGTGGTGC 848
DB 3659 CAGCTGCTCTCAAGGGGAGCGATGTGATAGCTGAGACCTGAGACAGTGTGGTGC 3718
QY 849 AGGTGGGTGGGTGACTACTATTTGCAATTTATGCAATGATCAAGACAGACAGACCTTCT 908
DB 3719 AGGTGGGCGTGGGTGATTAATTTGCAATTTATGCAATGATCAAGACAGACAGACCTTCT 3778
QY 909 CCGGATTTCTGAGTACTCCGACTGAGACAGCTCCCGAGCTTGTCTTACTGCGCACTGC 968
DB 3779 CTGATTTCTGCTTACTTCTGACTGAGACAGCTCCCGAGCTTGTCTTAAACAGCTGA 3838
QY 969 AAGTGAAGTCAATGCTCTCACTCTAGAGAGAGGTGAGCTGACCAACAGGCTATCC 1028
DB 3839 ACCGAGAGCTGGACACTTCTCTCAGAGAGGTGAGCTGACCAACCGGCGCAAGCTATC 3898
QY 1029 AGGAGGGCTGGCCCCCTGGAATATTTGTAATGACTAGGAGGTGGGTGAGACACTTC 1088
DB 3899 CAGGAGGGCTGGCCCCCTGGAATATTTGTAATGACTAGGAGAGAGAGAG-CAACTTCC 3957
QY 1089 CTTCTCTGCTGCTGAGCAAGAAATGGGAAACAGTGTCTCTCCGATCAGTCTGCGAGATG 1148
DB 3958 AGTCCCACTCTGCGCAATGATGAGACA--GGCTGTCTGAGTCA---AGACAGGTG 4011
QY 1149 GGGCAGAGGCTGATTTCTGCCCAAGACAGAGAGTGTCTGTGCTGCGAGTGAAT 1208
DB 4012 GAGCAGTGGCTGGGTTCTGCGCAGACTTGA--GAATGCAATGAGCTGCGAGCTGTGGT 4070
QY 1209 CCCCCAGTTCCTGTGTCAGAGAGCCCAAGGTGGGTGCTCTCTTCTGTGCTCTG--- 1265
DB 4071 C-----CTGGCCAGAGACTCCAGGTGGAGTGTCCATTCCTAGTCTGTGTCC 4119
QY 1266 CTTCTGTGATCTCTCCCAACCCCTCTCTGCTCTGCGGGCGCGCTTTTCTCAGAGATCA 1325
DB 4120 CTTCTAGTCTCTGACTCTCAATCTCTGTGCTCTCCAGGCGAGGCTTTTCTCAGAGATCA 4179
QY 1326 CTCAATAACCTAAGAACCTCATTAATAAAAAAAAAAAAAA 1366
DB 4180 CTTAATAACCTAATAATCTCTCCAAAAAATAAAAAAAAAA 4220

```

RESULT 10
 AP001003 173038 bp DNA linear HTG 30-MAY-2000
 LOCUS Homo sapiens chromosome 11 clone RP11-680C5 map 11q23, WORKING
 DEFINITION DRAFT SEQUENCE, 28 unordered pieces.
 ACCESSION AP001003
 VERSION AP001003.2 GI:8117674
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
 SOURCE Homo sapiens DNA, clone:RP11-680C5.
 ORGANISM Homo sapiens
 Buxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 173038)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 173,038 genomic DNA of 11q23
 TITLE Published Only in Database (2000)
 JOURNAL 2 (bases 1 to 173038)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 TITLE Submitted (05-JUN-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);


```

/misc_feature /note="assembly_fragment"
110123.118116
/misc_feature /note="assembly_fragment"
118217.126464
/misc_feature /note="assembly_fragment"
126565.132136
/misc_feature /note="assembly_fragment"
132237.135290
/misc_feature /note="assembly_fragment"
135391.140229
/misc_feature /note="assembly_fragment"
140330.143896
/misc_feature /note="assembly_fragment"
143997.147572
/misc_feature /note="assembly_fragment"
147673.150446
/misc_feature /note="assembly_fragment clone_end:T7 vector_side:right"
150547.153715
/misc_feature /note="assembly_fragment"
153816.157273
/misc_feature /note="assembly_fragment"
157374.159817
/misc_feature /note="assembly_fragment"
159918.161743
/misc_feature /note="assembly_fragment"
161844.163684
/misc_feature /note="assembly_fragment"
163785.165216
/misc_feature /note="assembly_fragment clone_end:SP6 vector_side:left"
165317.166657
/misc_feature /note="assembly_fragment"
166758.168307
/misc_feature /note="assembly_fragment"

Query Match 50.4%; Score 694; DB 2; Length 173038;
Best Local Similarity 98.8%; Pred. No. 4e-110;
Matches 741; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

QY 604 CCGCGTGTGTGAACGACGAGGACATTAACGACGCGCTCAACCGCAAGTTCACTGCA 663
Db 166758 CCGGTGTGTGTGAACGAGCA-GGACATTAAGACGCGCTCAACCGCAAGTTCACTGCA 166816
QY 664 GGTGCTGTGGGTCTAATCTTGGCGTCAATGCAACGCTTCAACGCGGCGACCTGCACTT 723
Db 166817 GGTGCTGTGGGTCTAATCTTGGCGTCAATGCAACGCTTCAACGCGGCGACCTGCACTT 166876
QY 724 TGATCTGTGAAGATGCGAATCATTCCTTTCTTCAGTTTTC-GGGGGGTGCGC 782
Db 166877 TGATCTGTGAAGATGCGAATCATTCCTTTCTTCAGTTTTCGGGGGGGTGCGC 166936
QY 783 CCAAGCCAGCTCTCGTCTGGGGGGGGGCAATGTAAGGCTGAGAGCTGAGAGCCAAAGTGT 842
Db 166937 CCAAGCCAGCTCTCGTCTGGGGGGGGGCAATGTAAGGCTGAGAGCTGAGAGCCAAAGTGT 166996
QY 843 GGGGCGAGGTGGGTGTGGGTGAATTAATGCAATGTAAGGCTGAGAGCTGAGAGCCAAAGTGT 902
Db 166997 GGGGCGAGGTGGGTGTGGGTGAATTAATGCAATGTAAGGCTGAGAGCTGAGAGCCAAAGTGT 167056
QY 903 CTTTCTCGGATTTCTGGTGAATCTCCAGTGGCAGAGCTCCCAAGTCTTTGCTTAGTGCC 962
Db 167057 CTTTCTCGGATTTCTGGTGAATCTCCAGTGGCAGAGCTCCCAAGTCTTTGCTTAGTGCC 167116
QY 963 CACTGCAAGTGAAGTCAATGCTCTCACTCTCAAGAGAGGGGTGAGGCTGACCAACAGG 1022
Db 167117 CACTGCAAGTGAAGTCAATGCTCTCACTCTCAAGAGAGGGGTGAGGCTGACCAACAGG 167176
QY 1023 TCAAT-CCAGAGAGGCTGAGCCGCC-TGGAATATTGTGAATGACTGAGAGGTGGGTGGA 1080
Db 167177 TCAATCCAGAGAGGCTGAGCCGCCCTTGGAATATTGTGACTGAGAGAGGTGGGTGGA 167236
QY 1081 GCACTCTCGTCTGCTGCTGAGCAAGATGGAGACAGTGTCTGTCGATCAAGTCTG 1140
Db 167237 GCACTCTCGTCTGCTGCTGAGCAAGATGGAGACAGTGTCTGTCGATCAAGTCTG 167296

```

```

QY 1141 GCAGCATTGGGACAGTGTGATTTCTGCCCCAAGACCAAGAGAGTGTCTGTGCGCA 1200
Db 167297 GCAGCATTGGGACAGTGTGATTTCTGCCCCAAGACCAAGAGAGTGTCTGTGCGCA 167356
QY 1201 GTGTAAGTCCCGCAGTGTCTGTGATCCAGAGCCCAAGGTGGGGTGTCTCTTCTGTGTC 1260
Db 167357 GTGTAAGTCCCGCAGTGTCTGTGATCCAGAGCCCAAGGTGGGGTGTCTCTTCTGTGTC 167416
QY 1261 CTGTGCTTCTGTGATCTTCCCAACCCCTCTGCTCTGAGGCGGCGCTTTTCTGAGA 1320
Db 167417 CTGTGCTTCTGTGATCTTCCCAACCCCTCTGCTCTGAGGCGGCGCTTTTCTGAGA 167476
QY 1321 GATCAGCATTAACCTTAAGACCTTCATA 1350
Db 167477 GATCAGCATTAACCTTAAGACCTTCATA 167506

RESULT 11
BC025174 1234 bp mRNA linear ROD 07-AUG-2002
LOCUS BC025174
DEFINITION Mus musculus, clone MGC:36714 IMAGE:3978387, mRNA, complete cds.
ACCESSION BC025174
VERSION BC025174.1 GI:19263583
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1234)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: ggaabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gumartine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNI.ac: http://image.llnl.gov
Series: IRAC Plate: 61 Row: f Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORP
analysis.
FEATURES
source
location/Qualifiers
1.1234
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:36714 IMAGE:3978387"
/tissue_type="Mammary tumor. Metallochromien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
112.843
/codon_start=1
/product="unknown (protein for MGC:36714)"
/protein_id="AA25174.1"
/db_xref="GI:19263584"

```

CDS

are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1
20728 20627 contig of 20627 bp in length
31918 31817 contig of 11090 bp in length
41232 49745 contig of 8514 bp in length
49846 59457 contig of 9612 bp in length
59558 66804 contig of 7247 bp in length
66805 73970 contig of 7066 bp in length
74071 81007 contig of 6937 bp in length
81108 87320 contig of 6213 bp in length
87421 93250 contig of 5830 bp in length
93251 97965 contig of 4615 bp in length
98066 102941 contig of 4876 bp in length
103042 109017 contig of 5976 bp in length
109118 113391 contig of 4274 bp in length
113492 118783 contig of 5292 bp in length
118884 124241 contig of 5358 bp in length
124342 129766 contig of 5425 bp in length
129867 135473 contig of 5607 bp in length
135574 140585 contig of 5012 bp in length
140686 144027 contig of 3342 bp in length
144128 147923 contig of 3796 bp in length
148024 152461 contig of 4438 bp in length
152562 155444 contig of 2883 bp in length
155545 158721 contig of 3177 bp in length
158822 162903 contig of 4082 bp in length
163004 166121 contig of 3118 bp in length
166222 169851 contig of 3630 bp in length
169952 172048 contig of 2097 bp in length
172149 175332 contig of 3184 bp in length
175433 177914 contig of 2482 bp in length
178015 179875 contig of 1861 bp in length
179976 181604 contig of 1629 bp in length
181705 183951 contig of 2247 bp in length
184052 185927 contig of 1876 bp in length
186028 187870 contig of 1843 bp in length
187971 189591 contig of 1621 bp in length
189692 191169 contig of 1478 bp in length
191270 192527 contig of 1258 bp in length
192628 194288 contig of 1661 bp in length
194389 195658 contig of 1270 bp in length
195759 197707 contig of 1949 bp in length
197808 198902 contig of 1095 bp in length
Sequence updated (26-May-2000).
NOTE: This is a 'working draft' sequence. It currently
consists of 42 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
20627 20627 contig of 20627 bp in length
20728 20727 gap of 100 bp
31818 31817 contig of 11090 bp in length
31918 41131 contig of 9214 bp in length
41132 41231 gap of 100 bp
41233 49745 contig of 8514 bp in length
49746 49845 gap of 100 bp
49846 59457 contig of 9612 bp in length
59458 59557 gap of 100 bp
59558 66804 contig of 7247 bp in length
66805 66904 gap of 100 bp
66905 73970 contig of 7066 bp in length
73971 74070 gap of 100 bp
74071 81007 contig of 6937 bp in length
81008 81107 gap of 100 bp
81108 87320 contig of 6213 bp in length
87321 87420 gap of 100 bp
87421 93250 contig of 5830 bp in length

```

FEATURES

Source

```

* 93251 93350: gap of 100 bp
* 93351 97965: contig of 4615 bp in length
* 97966 98065: gap of 100 bp
* 98066 102941: contig of 4876 bp in length
* 102942 103041: gap of 100 bp
* 103042 109017: contig of 5976 bp in length
* 109018 109117: gap of 100 bp
* 109118 113391: contig of 4274 bp in length
* 113392 113491: gap of 100 bp
* 113492 118783: contig of 5292 bp in length
* 118784 118883: gap of 100 bp
* 118884 124241: contig of 5358 bp in length
* 124242 124341: gap of 100 bp
* 124342 129766: contig of 5425 bp in length
* 129767 129866: gap of 100 bp
* 129867 135473: contig of 5607 bp in length
* 135474 135573: gap of 100 bp
* 135574 140585: contig of 5012 bp in length
* 140586 140685: gap of 100 bp
* 140686 144027: contig of 3342 bp in length
* 144028 144127: gap of 100 bp
* 144128 147923: contig of 3796 bp in length
* 147924 148023: gap of 100 bp
* 148024 152461: contig of 4438 bp in length
* 152462 152561: gap of 100 bp
* 152562 155444: contig of 2883 bp in length
* 155445 155544: gap of 100 bp
* 155545 158721: contig of 3177 bp in length
* 158722 158821: gap of 100 bp
* 158822 162903: contig of 4082 bp in length
* 162904 163003: gap of 100 bp
* 163004 166121: contig of 3118 bp in length
* 166122 166221: gap of 100 bp
* 166222 169851: contig of 3630 bp in length
* 169852 169951: gap of 100 bp
* 169952 172048: contig of 2097 bp in length
* 172049 172148: gap of 100 bp
* 172149 175332: contig of 3184 bp in length
* 175333 175432: gap of 100 bp
* 175433 177914: contig of 2482 bp in length
* 177915 178014: gap of 100 bp
* 178015 179875: contig of 1861 bp in length
* 179876 179975: gap of 100 bp
* 179976 181604: contig of 1629 bp in length
* 181605 181704: gap of 100 bp
* 181705 183951: contig of 2247 bp in length
* 183952 184051: gap of 100 bp
* 184052 185927: contig of 1876 bp in length
* 185928 186027: gap of 100 bp
* 186028 187870: contig of 1843 bp in length
* 187871 187970: gap of 100 bp
* 187971 189591: contig of 1621 bp in length
* 189592 189691: gap of 100 bp
* 189692 191169: contig of 1478 bp in length
* 191170 191269: gap of 100 bp
* 191270 192527: contig of 1258 bp in length
* 192528 192627: gap of 100 bp
* 192628 194288: contig of 1661 bp in length
* 194289 194388: gap of 100 bp
* 194389 195658: contig of 1270 bp in length
* 195659 195758: gap of 100 bp
* 195759 197707: contig of 1949 bp in length
* 197708 197807: gap of 100 bp
* 197808 198902: contig of 1095 bp in length.
Location/Qualifiers
1. 198902
/organism="Homo sapiens"
/db xref="taxon:9606"
/chromosome="11"
/map="11q23"
/clone="RP11-657A24"
1. 20627
/note="assembly_fragment"
misc_feature

```


* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 1137: contig of 1137 bp in length
 * 1138 1237: gap of unknown length
 * 1238 2720: contig of 1483 bp in length
 * 2721 2820: gap of unknown length
 * 2821 3984: contig of 1164 bp in length
 * 3985 4084: gap of unknown length
 * 4085 5158: contig of 1074 bp in length
 * 5159 5258: gap of unknown length
 * 5259 6306: contig of 1048 bp in length
 * 6307 6406: gap of unknown length
 * 6407 7523: contig of 1117 bp in length
 * 7524 7623: gap of unknown length
 * 7624 8631: contig of 1008 bp in length
 * 8632 8731: gap of unknown length
 * 8732 9806: contig of 1075 bp in length
 * 9807 9906: gap of unknown length
 * 9907 11155: contig of 1249 bp in length
 * 11156 11255: gap of unknown length
 * 11256 12568: contig of 1313 bp in length
 * 12569 12668: gap of unknown length
 * 12669 13704: contig of 1036 bp in length
 * 13705 13804: gap of unknown length
 * 13805 14931: contig of 1127 bp in length
 * 14932 15031: gap of unknown length
 * 15032 16084: contig of 1053 bp in length
 * 16085 16184: gap of unknown length
 * 16185 17858: contig of 1674 bp in length
 * 17859 19035: gap of unknown length
 * 19036 19135: contig of 1077 bp in length
 * 19136 20291: gap of unknown length
 * 20292 20391: gap of unknown length
 * 20392 21643: contig of 1252 bp in length
 * 21644 21743: gap of unknown length
 * 21744 22846: contig of 1103 bp in length
 * 22847 22946: gap of unknown length
 * 22947 24684: contig of 1738 bp in length
 * 24685 24784: gap of unknown length
 * 24785 26405: contig of 1621 bp in length
 * 26406 26505: gap of unknown length
 * 26506 27716: contig of 1211 bp in length
 * 27717 27817: gap of unknown length
 * 27818 28876: contig of 1060 bp in length
 * 28877 28976: gap of unknown length
 * 28977 30019: contig of 1043 bp in length
 * 30020 30119: gap of unknown length
 * 30120 31550: contig of 1431 bp in length
 * 31551 31650: gap of unknown length
 * 31651 33128: contig of 1478 bp in length
 * 33129 33228: gap of unknown length
 * 33229 34763: contig of 1535 bp in length
 * 34764 34863: gap of unknown length
 * 34864 36404: contig of 1541 bp in length
 * 36405 36504: gap of unknown length
 * 36505 37957: contig of 1453 bp in length
 * 37958 38057: gap of unknown length
 * 38058 40293: contig of 2236 bp in length
 * 40294 40393: gap of unknown length
 * 40394 41698: contig of 1305 bp in length
 * 41699 41798: gap of unknown length
 * 41799 43155: contig of 1357 bp in length
 * 43156 43255: gap of unknown length
 * 43256 45193: contig of 1938 bp in length
 * 45194 45293: gap of unknown length
 * 45294 47217: contig of 1924 bp in length
 * 47218 47317: gap of unknown length
 * 47319 49231: contig of 1914 bp in length
 * 49232 49331: gap of unknown length

* 49332 51205: contig of 1874 bp in length
 * 51206 51305: gap of unknown length
 * 51306 53117: contig of 1812 bp in length
 * 53118 53217: gap of unknown length
 * 53218 55542: contig of 2325 bp in length
 * 55543 55642: gap of unknown length
 * 55643 57451: contig of 1809 bp in length
 * 57452 57552: gap of unknown length
 * 57553 60851: contig of 3300 bp in length
 * 60852 60951: gap of unknown length
 * 60952 62718: contig of 1767 bp in length
 * 62719 62818: gap of unknown length
 * 62819 64976: contig of 2158 bp in length
 * 64977 65076: gap of unknown length
 * 65077 67855: contig of 2779 bp in length
 * 67856 67955: gap of unknown length
 * 67956 71313: contig of 3358 bp in length
 * 71314 71413: gap of unknown length
 * 71414 73659: contig of 2246 bp in length
 * 73660 73759: gap of unknown length
 * 73760 75549: contig of 1780 bp in length
 * 75550 75649: gap of unknown length
 * 75650 79106: contig of 3457 bp in length
 * 79107 79206: gap of unknown length
 * 79207 83551: contig of 4345 bp in length
 * 83552 83651: gap of unknown length
 * 83652 86053: contig of 2402 bp in length
 * 86054 86153: gap of unknown length
 * 86154 88659: contig of 2506 bp in length
 * 88660 90078: gap of unknown length
 * 90079 90178: contig of 1319 bp in length
 * 90179 92201: contig of 2023 bp in length
 * 92202 92302: gap of unknown length
 * 92303 94877: contig of 2576 bp in length
 * 94878 94977: gap of unknown length
 * 94978 96946: contig of 1969 bp in length
 * 96947 97046: gap of unknown length
 * 97047 99897: contig of 2851 bp in length

Query Match 37.3%; Score 513.4; DB 2; Length 173657;
 Best Local Similarity 78.9%; Pred. No. 5.2e-79;
 Matches 717; Conservative 0; Mismatches 171; Indels 21; Gaps 8;

QY 438 CGGAGCTGCGGAGCTGAGGAGGAGCCCGGGGCGCGAGAGAGCGGAGCCCGGGGCG 497
 DB 43765 CAGGACTACCTGGGCGACGCTGGGAGAGCCCGGGGCGCGTGAAGAGCGAGCCTGTGGGG 43824
 QY 438 CCACCGGCGCTGCGGAGAGTCTCGTGCCTCCCGCATCCGCTTCAAGCGCAAGCGCT 557
 DB 43825 CTAATGGGGCTGCGGGGAGTGTCTGGTGTCCCGCATCAAGCTTCACTGATGCGCAAGCAT 43884
 QY 558 CCGAAGCCCGGCTGCTCGCGGCTTGAAGCACCTTGGCTTCAACCGCGTGTGTGA 617
 DB 43885 CAGAGAGCGGGTACTCGCGAGCGACACACCTTCAACCGTGTGTGTCTCA 43944
 QY 618 ACGAGAGGAGCATTACGACGCGGACCGGAGAGTTACCTGCGAGGCTGTGGGCT 677
 DB 43945 ATGAGAGGAGCATTACGATGCTACCTACCGGAGAGTTACCTGCGAGGCTGTGGGCT 44004
 QY 678 ACTACTTGGCGCTGATGCAACCGTCTACCGGGGAGCGCTGAGTTGATGTGAAGA 737
 DB 44005 ACTACTTGGCTGATGCAACCGTCTACCGGGGAGCGCTGAGTTGATGTGAAGA 44064
 QY 738 ATGGGAGATTCATTCCTTTCTTCAGTTTTCGGGGGGTGGCCCAAGCCAGCTGCG 797
 DB 44065 ATGGGCAATTCATTCCTTTCTTCAGTTTTCGGGGGGTGGCCCAAGCCAGCTGCG 44124
 QY 798 TCTCGGGGGGGGCATATGAGAGGCTGAGAGCTTGAAGCAACAGTGTGGGAGGAGGGG 857
 DB 44125 TCTCAGGGGGGGTGCATATGAGAGGCTTGAAGCAACAGTGTGGGAGGAGGGG 44184
 QY 858 TGGGTGACTACATTGGGCATCTATGCGACGATCAGAGACAGACCTTCTCCGGATTTC 917

PR 10-JUN-1998; 98US-0088742.
 PR 10-NOV-1998; 98US-0107783.
 XX (GENETH) GENENTECH INC.
 PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
 PI Roy MA, Wood WJ;
 XX WPI, 1999-430385/36.
 DR P-PSDP; AAY06481.
 XX
 PT Antibody against proteins expressed in neoplastic cells, useful for
 PT tumor diagnosis and treatment
 XX
 PS Example 1; Fig 9; 162pp; English.
 XX
 CC This is the nucleotide sequence of cDNA clone DNA40592 (ATCC 209492)
 CC coding for human PRO344 (UN0303) (see AAY06481). The clone was
 CC isolated from a foetal kidney library. Amplification of DNA40592
 CC was observed in primary lung tumours and in primary colon tumours,
 CC suggesting an association with tumour formation or growth.
 CC Antagonists (e.g. antibodies) directed against PRO344 may have
 CC utility in cancer therapy. The invention identifies 14 genes (see
 CC AAX87254-67) that are amplified in the genome of tumour cells. Such
 CC amplification is expected to be associated with overexpression of
 CC the gene product and to contribute to tumorigenesis. The encoded
 CC proteins (see AAY06477-90) may be useful targets for the diagnosis
 CC and/or treatment (including prevention) of certain cancers, and may
 CC act as predictors of the prognosis of tumour treatment.
 XX
 SQ Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;
 Query Match 100.0%; Score 1377; DB 20; Length 1377;
 Best Local Similarity 100.0%; Pzed. No. 1.7e-244;
 Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 541 CTTACAGCCGCAAGCGCTCCGAGAGCGGCTGCTCCGCGTCTGAGCAGCACCCTTGCCCTT 600
 Qy 601 CGACCGCGCTGCTGTTGAAGAGCAGAGGACATTACAGACCGCGTCAACCGGCAATTACCTG 660
 Db 601 CGACCGCGCTGCTGTTGAAGAGCAGAGGACATTACAGACCGCGTCAACCGGCAATTACCTG 660
 Qy 661 CCAGGTGCTTGGGGGCTTACTACTTCCGCTCCATGCGACCGTCTACCGGGGAGCCTTGA 720
 Db 661 CCAGGTGCTTGGGGGCTTACTACTTCCGCTCCATGCGACCGTCTACCGGGGAGCCTTGA 720
 Qy 721 GTTTGATCTGTGAAGATGCGAATTCATTGCTCTTTTCTTCAGTTTTCGGGGGGTG 780
 Db 721 GTTTGATCTGTGAAGATGCGAATTCATTGCTCTTTTCTTCAGTTTTCGGGGGGTG 780
 Qy 781 GCCCAAGCAGGCTGCTCGGGGGGGGGCCATGTTGAGGCTGAGGCTCAGAGCAACAAT 840
 Db 781 GCCCAAGCAGGCTGCTCGGGGGGGGGCCATGTTGAGGCTGAGGCTCAGAGCAACAAT 840
 Qy 841 GTGGGTGCAAGTGGGTGTGGTGAATTAATGGAATCTATGCGACATCAAGACAGACAG 900
 Db 841 GTGGGTGCAAGTGGGTGTGGTGAATTAATGGAATCTATGCGACATCAAGACAGACAG 900
 Qy 901 CACCTTCTCCGATTTTCTGTGTACTCGACCTGCGACAGCTCCCACTTTTGTAGTG 960
 Db 901 CACCTTCTCCGATTTTCTGTGTACTCGACCTGCGACAGCTCCCACTTTTGTAGTG 960
 Qy 961 CCCACTGCAAGTGAAGCTCATGCTCTCACTCTAGAGGAGGTGAGAGCTGACAAACA 1020
 Db 961 CCCACTGCAAGTGAAGCTCATGCTCTCACTCTAGAGGAGGTGAGAGCTGACAAACA 1020
 Qy 1021 GGTATTCAGAGAGGCTGCGCCCTCGAATAATTTGTAATGACTAGGAGTGGGTAGA 1080
 Db 1021 GGTATTCAGAGAGGCTGCGCCCTCGAATAATTTGTAATGACTAGGAGTGGGTAGA 1080
 Qy 1081 GCACTCTCCGCTCTGCTGCTGCGAAGATGGAACAGTGGCTGCTGATCAGGTCTG 1140
 Db 1081 GCACTCTCCGCTCTGCTGCTGCGAAGATGGAACAGTGGCTGCTGATCAGGTCTG 1140
 Qy 1141 GCAGCATGGGGCAGTGGCTGATTTCTGCCCCAAGACAGAGAGTGTGCTGAGCA 1200
 Db 1141 GCAGCATGGGGCAGTGGCTGATTTCTGCCCCAAGACAGAGAGTGTGCTGAGCA 1200
 Qy 1201 GTGTAAATGCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db 1201 GTGTAAATGCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Qy 1261 CTCTGCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 Db 1261 CTCTGCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 Qy 1321 GATCACTCAATTAACCTTAAGAACTCTCAATTAAGAACTCTCAATTAAGAACTCTCA 1377
 Db 1321 GATCACTCAATTAACCTTAAGAACTCTCAATTAAGAACTCTCAATTAAGAACTCTCA 1377
 RESULT 2
 AAX80052
 ID AAX80052 standard; cDNA, 1377 BP.
 XX
 AC AAX80052;
 XX
 DT 12-AUG-1999 (first entry)
 XX
 DE Human PRO344 nucleotide sequence.
 XX
 DE Human; PRO protein; tumour necrosis factor family; TNF; cytokine;
 XX
 KW secreted protein; transmembrane protein; inflammation disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W09928462-A2.
 XX

QY	121	CCAACTGGAGGGTCCCGAGTAAAGGACGCCCCGGAAGAGAGCCATGGGAGACCCGAGAGG	180
Db	121	CCAACTGGAGGGTCCCGAGTAAAGGAGCCCCGGAAGAGAGCCATGGGAGACCCGAGAGG	180
QY	181	GGAGACTGGAGAGGACCCCGCGCTCCGGGCTCCCGGTGCAGCGCTATGAGGACCACTCT	240
Db	181	GGGACTGGAGAGGACCCCGGCGCTCCGGGCTCCCGGTGCAGCGCTATGAGGCCACTCTCT	240
QY	241	GGTCTGTGTGTCTCTGGGCTGTGGCGGCTGGCTGTGCCCCCTGACGTAGCAACAAGATCCC	300
Db	241	GGTCTGTGTGTCTCTGGGCTGTGGCGGCTGGCTGTGCCCCCTGACGTAGCAACAAGATCCC	300
QY	301	CAGCCTCTGACCCGGGAGCACCCGGCTTCAAGGACAGCGGGGCAACAATGACAGCAGG	360
Db	301	CAGCCTCTGACCCCGGGAGCACCCGGCTTCAAGGACAGCGGGGCAACAATGACAGCAGG	360
QY	361	CTTTCCTGGGCTCCGATGTGCTGCGACGGCCGCAAGGAGCGCCCGGGGCTCCGGAGAGAA	420
Db	361	CTTTCCTGGGCTCCGATGTGCTGCGACGGCCGCAAGGAGCGCCCGGGGCTCCGGAGAGAA	420
QY	421	AGGAGAGGGGCGGAGAGCGGGGACTGGCCGGGACCTGAGAGGGGACCCCGGGCCGCGAGAG	480
Db	421	AGGAGAGGGGCGGAGAGCGGGGACTGGCCGGGACCTGAGAGGGGACCCCGGGCCGCGAGAG	480
QY	481	GGCGGAGACCCGCGGGGCCCAACCGGGCTTGCCTGGGAGAGTGTCTGGGCTCTCGCGAGTCCG	540
Db	481	GGCGGAGACCCGCGGGGCCCAACCGGGCTTGCCTGGGAGAGTGTCTGGGCTCTCGCGAGTCCG	540
QY	541	CTTCAAGGCGCAAGCGCTCCGAGAGCGCGGGTGTCTTCCCGGTCTGACGACCTTGGCCCTT	600
Db	541	CTTCAAGGCGCAAGCGCTCCGAGAGCGCGGGTGTCTTCCCGGTCTGACGACCTTGGCCCTT	600
QY	601	CGACCGGCTGTGTGTGAACGAGACAGGACATTAACGACGCGCTCAACCGGCAAGTTCACTG	660
Db	601	CGACCGGCTGTGTGTGAACGAGACAGGACATTAACGAGCGCTCAACCGGCAAGTTCACTG	660
QY	661	CGAGGTGCTGGGGGTCTACTACTTTCGGCTCAATGCGCACGCTCTACCGGGCCAGCTGCA	720
Db	661	CGAGGTGCTGGGGGTCTACTACTTTCGGCTCAATGCGCACGCTCTACCGGGCCAGCTGCA	720
QY	721	GTTTGAATCTGTGAAGATAGCGAATTCATGTCCTTCTTCAAGTTTTCGGGGGGTGG	780
Db	721	GTTTGAATCTGTGAAGATAGCGAATTCATGTCCTTCTTCAAGTTTTCGGGGGGTGG	780
QY	781	GGCCAAGCCAGCTCGCTCTCGGGGGGGGCGATGTGAAGCTGAGGCTTGAGACCAAGT	840
Db	781	GGCCAAGCCAGCTCGCTCTCGGGGGGGGCGATGTGAAGCTGAGGCTTGAGACCAAGT	840
QY	841	GTGGGTGCAAGTGGGTGTGATGATCAATTGGGATCTAATGCCAGCATCAAGACACAG	900
Db	841	GTGGGTGCAAGTGGGTGTGATGATCAATTGGGATCTAATGCCAGCATCAAGACACAG	900
QY	901	CACCTTCTCCGATTTCTGTGTGTACTCCGACTGACAGCTCCCAAGTCTTTGCTTAAGT	960
Db	901	CACCTTCTCCGATTTCTGTGTGTACTCCGACTGACAGCTCCCAAGTCTTTGCTTAAGT	960
QY	961	CCCACTGCAAAATGAGCTCATGCTCTCACTCTTGAAGAGGGGTGTGAGGCTTGACACA	1020
Db	961	CCCACTGCAAAATGAGCTCATGCTCTCACTCTTGAAGAGGGGTGTGAGGCTTGACACA	1020
QY	1021	GGTCAATCCAGAGGGGCTGCCCCCTGTGATATTGTGAATGACTAAGAGAGTGGGTTAGA	1080
Db	1021	GGTCAATCCAGAGGGGCTGCCCCCTGTGATATTGTGAATGACTAAGAGAGTGGGTTAGA	1080
QY	1081	GCACCTCTCCGTCTCTCTGTGTGAGCAAGAAATGGGAAAGTGGCTGTGTGCTGGAG	1140
Db	1081	GCACCTCTCCGTCTCTCTGTGTGAGCAAGAAATGGGAAAGTGGCTGTGTGCTGGAG	1140
QY	1141	GCAGCAGTGGGAGAGTGTCTGAATTTCTGCCAAGACCAAGAGAGTGTGTGTGCTGGAG	1200
Db	1141	GCAGCAGTGGGAGAGTGTCTGAATTTCTGCCAAGACCAAGAGAGTGTGTGTGCTGGAG	1200


```

Db      1201  GTGTAAAGTCCCGCCAGTTGCTGTGTCAGAGAGCCCAAGGAGGAGTCTCTCTTCCTGGTCTC
Oy      1261  CTCTGCTCTCTCTGATGATCTCTCCCAAGCCCTCTGCTCTGGAGGAGCCCTTTTCTCAGA
Db      1261  CTCTGCTCTCTCTGATGATCTCTCCCAAGCCCTCTGCTCTGGAGGAGCCCTTTTCTCAGA
Oy      1321  GATCACTCAATTAACCTTAAGAACCTCATTAATAAAAAAAAAAAAAAAAAAAAAA 1377
Db      1321  GATCACTCAATTAACCTTAAGAACCTCATTAATAAAAAAAAAAAAAAAAAAAAAA 1377

```

RESULT 4

AAA46907

ID AAA46907 standard; cDNA; 1377 BP.

AC AAA46907;

XX 03-OCT-2000 (first entry)

DE cDNA encoding novel polypeptide PRO344.

XX PRO201; PRO292; PRO327; PRO1265; PRO344; PRO343; PRO347; PRO357;

XX PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell; tumourigenesis; cancer; neoplastic cell growth; cell proliferation; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 227..958

FT /tag= a

XX WO200037640-A2.

XX 29-JUN-2000.

XX 16-DEC-1999; 99WO-US30095.

XX 22-DEC-1998; 98US-0113296.

XX 08-MAR-1999; 99WO-US05028.

XX 02-JUN-1999; 99WO-US12252.

XX 01-SEP-1999; 99WO-US20111.

XX 15-SEP-1999; 99WO-US21090.

XX 30-NOV-1999; 99WO-US28313.

XX 01-DEC-1999; 99WO-US28409.

XX 02-DEC-1999; 99WO-US28301.

XX (GETH) GENENTECH INC.

XX Borstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;

XX Wood WI;

XX WPI; 2000-452188/39.

XX P-PSDB; AAY93688.

XX New anti-polypeptide antibody useful in the treatment and diagnosis of

XX neoplastic cell growth and proliferation -

XX Claim 50; Fig 9; 220pp; English.

XX The present sequence encodes a novel human polypeptide. The

XX specification describes novel polypeptides designated PRO201, PRO292,

XX PRO327, PRO1265, PRO344, PRO343, PRO347, PRO357, PRO1017,

XX PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in

XX the genome of tumour cells. The polypeptides are believed to contribute

XX to tumourigenesis. The polypeptides are useful target for the

XX identification of certain cancers, and may act as predictors of the

XX prognosis of tumour treatment. Antibodies against these polypeptides

XX are useful in the treatment and diagnosis of neoplastic cell growth

XX and proliferation in mammals.

XX Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;

Query Match 100.0%; Score 1377; DB 21; Length 1377;
 Best Local Similarity 100.0%; Pred. No. 1.7e-244;
 Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      1  GACTAGTTCTCTTGAAGTCTTGGAGAGAGAAAGCGAGAGCCGAGAGCGAAGCCAGAGAC 60
Db      1  GACTAGTTCTCTTGAAGTCTTGGAGAGAGAAAGCGAGAGCCGAGAGCGAAGCCAGAGAC 60
Oy      61  TGGGGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db      61  TGGGGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Oy      121  CCAACTGAAGAGTCCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db      121  CCAACTGAAGAGTCCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Oy      181  GGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db      181  GGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Oy      241  CGTCTGCTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db      241  CGTCTGCTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Oy      301  CAGCTCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db      301  CAGCTCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Oy      361  CTTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db      361  CTTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Oy      421  AGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db      421  AGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Oy      481  GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      481  GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Oy      541  CTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db      541  CTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Oy      601  CGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      601  CGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Oy      661  CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      661  CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Oy      721  GTTGTATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      721  GTTGTATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Oy      781  GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db      781  GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Oy      841  GTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db      841  GTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Oy      901  CACCTTCTGCGAGATTTCTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db      901  CACCTTCTGCGAGATTTCTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Oy      961  CCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db      961  CCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020

```


Db 841 GTGGGTGACAGTGGGTGTGGTGTACCTACATTTGGCATCTATGCCAGCATCAAGACAGACAG 900
 Qy 901 CACCTTCTCCGGATTTCTGTGTATCTCCGACTGGACAGCTCCCAAGTCTTTGCTTAGTG 960
 Db 901 CACCTTCTCCGGATTTCTGTGTATCTCCGACTGGACAGCTCCCAAGTCTTTGCTTAGTG 960
 Qy 961 CCCACTGCAAGTACGTCTCATGTCTCTCATCTCTGAAAGAGAGGTGTAGAGCTGACAA 1020
 Db 961 CCCACTGCAAGTACGTCTCATGTCTCTCATCTCTGAAAGAGAGGTGTAGAGCTGACAA 1020
 Qy 1021 GGTTCATCAGAGAGGCTGGGCCCCCTGGAAATATTTGTGAATGACATGAGGAGGTGGGTAGA 1080
 Db 1021 GGTTCATCAGAGAGGCTGGGCCCCCTGGAAATATTTGTGAATGACATGAGGAGGTGGGTAGA 1080
 Qy 1081 GCACCTCTCCGCTCTGCTGTGGACAGAGATGGAAACAGTGTCTGTGCATCAGGTCTG 1140
 Db 1081 GCACCTCTCCGCTCTGCTGTGGACAGAGATGGAAACAGTGTCTGTGCATCAGGTCTG 1140
 Qy 1141 GCAGCATTGGGGGAGTGTGTGATTTTCTGCCCAAGACAGAGAGTGTGTGTCTGCGCA 1200
 Db 1141 GCAGCATTGGGGGAGTGTGTGATTTTCTGCCCAAGACAGAGAGTGTGTGTCTGCGCA 1200
 Qy 1201 GTGTAAAGTCCCCAGTGTCTCTGTGTCAGAGAGCCAGGTGGGTGCTCTTCCCTGGTC 1260
 Db 1201 GTGTAAAGTCCCCAGTGTCTCTGTGTCAGAGAGCCAGGTGGGTGCTCTTCCCTGGTC 1260
 Qy 1261 CTCTGCTCTCTGTGTGATCTCCCAACCCCTCTCTGTCTGTGGGCGCGCCCTTTTCTCAGA 1320
 Db 1261 CTCTGCTCTCTGTGTGATCTCCCAACCCCTCTCTGTCTGTGGGCGCGCCCTTTTCTCAGA 1320
 Qy 1321 GATCAGCTCATTAACCTAAGAACCTCTATTAACCTCTATTAACCTCTATTAACCTCTATTA 1377
 Db 1321 GATCAGCTCATTAACCTAAGAACCTCTATTAACCTCTATTAACCTCTATTAACCTCTATTA 1377
 RESULT 6
 AAS21424
 AAS21424 standard; cDNA; 1377 BP.
 AC AAS21424;
 XX 24-OCT-2001 (first entry)
 DE Human cDNA sequence encoding for PRO344 polypeptide.
 XX
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200140466-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 01-DEC-2000; 2000WO-US32678.
 XX
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 30-DEC-1999; 99WO-US30999.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.

PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 10-NOV-2000; 2000WO-US30873.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
 PI Gerritsen MB, Goddard A, Godowski PJ, Gunney AL, Sherwood S,
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI; 2001-408281/43.
 DR P-PSDB; AAIJ2352.
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 PS Claim 3; Fig 361; 813pp; English.
 XX
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX
 SQ Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;
 Query Match 100.0%; Score 1377; DB 22; Length 1377;
 Best Local Similarity 100.0%; Pred. No. 1,7e-244;
 Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GACTAGTCTCTTGTGAGTCTGTGAGAGAGAAAGCGGACCGGACGCAACACAGAC 60
 Db 1 GACTAGTCTCTTGTGAGTCTGTGAGAGAGAAAGCGGACCGGACGCAACACAGAC 60
 Qy 61 TGGGGTGAAGCGGACGAGGCGGCGCTGTGCGCGGGGAGAAAGCGGCGGTGAGACCA 120
 Db 61 TGGGGTGAAGCGGACGAGGCGGCGCTGTGCGCGGGGAGAAAGCGGCGGTGAGACCA 120
 Qy 121 CCAACTGAGAGGCTCGGAGTAGAGAGGCCCCGGAAGAGCCATCGGCGAGCGCGGAGG 180
 Db 121 CCAACTGAGAGGCTCGGAGTAGAGAGGCCCCGGAAGAGGCAATCGGCGAGCGCGGAGG 180
 Qy 181 GGAAGTCAAGAGAGACCCCGGCGTCCCGGTGACAGCGCTATGAGGCGCACTCT 240
 Db 181 GGAAGTCAAGAGAGACCCCGGCGTCCCGGTGACAGCGCTATGAGGCGCACTCT 240
 Qy 241 GGTCTCTGTCTCTGTGGGCTGTGGGCGCGGCTGCGCCCACTGAGCAACAAGATCCC 300

Db 1321 GATGACTCATATAACCTTAGAACCCCTCATTAACCAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 7
ID AAC58626 standard; cDNA; 1377 BP.
XX AAC58626;
XX
DT 29-JAN-2001 (first entry)
DE Human PRO344 protein UNQ303 encoding cDNA SEQ ID NO:240.

KW Human; immune related disease; diagnosis; antiinflammatory; cardiant;
KW dermatological; antlarthritis; antirheumatic; immunosuppressive;
KW haemostatic; antihypoid; antidiabetic; nootropic; neuroprotective;
KW antianaemic; hepatocytic; vitruide; antisoriatic; antiallergic;
KW antiaathmatic; systemic lupus erythematosus; rheumatoid arthritis;
KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
KW autoimmune chromocytopenia; immune-mediated renal disease;
KW demyelinating disease; hepatobiliary disease; Whipple's disease;
KW inflammatory bowel disease; gluten-sensitive enteropathy;
KW autoimmune disease; immune-mediated skin disease; allergic disease;
KW immunological disease; graft-versus-host-disease; ss.
KW graft rejection; graft-versus-host-disease; ss.
OS Homo sapiens.
XX
XX WO200053758-A2.
PN 14-SEP-2000.
XX PD
XX
XX 02-MAR-2000; 2000MO-US05841.
PF
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99US-0123618.
PR 12-MAR-1999; 99US-0123957.
PR 23-MAR-1999; 99US-0125775.
PR 12-APR-1999; 99US-0128849.
PR 20-APR-1999; 99WO-US08615.
PR 28-APR-1999; 99US-0131445.
PR 04-MAY-1999; 99US-0132377.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21050.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-OCT-1999; 99US-0162506.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000MO-US00219.
PR 06-JAN-2000; 2000MO-US00277.
PR 06-JAN-2000; 2000MO-US00376.
PR 11-FEB-2000; 2000MO-US03565.
PR 18-FEB-2000; 2000MO-US04341.

PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.

XX
XX
XX (GENT) GENENTECH INC.

PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W,
PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V,
PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M,

DR WPI: 2000-572271/53.
DR F-PSDB, AAB33461.

XX Sixty four PRO polypeptides, useful in the diagnosis and treatment of
PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -

XX
XX
XX Claim 23; Fig 95; 309pp; English.

XX The present invention describes sixty four human PRO proteins which can
CC be used in the treatment of immune related diseases. The human PRO
CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
CC treating and diagnosing immune related disorders. The disorders are
CC selected from systemic lupus erythematosus, rheumatoid arthritis,
CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis,
CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
CC immune-mediated renal diseases, demyelinating diseases of the central
CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,
CC autoimmune or immune-mediated skin diseases, allergic diseases,
CC immunological diseases of the lung, and transplantation associated
CC diseases including graft rejection and graft-versus-host-disease.
CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.

XX
XX
XX Sequence 1377 BP; 251 A; 423 C; 471 G; 232 T; 0 other;

Query Match 99.9%; Score 1375.4; DB 21; Length 1377;

Best Local Similarity 99.9%; Pred. No. 3.3e-244;

Matches 1376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 GACTAGTCTCTTGAAGTGTGAGAGAGAAAGCCGAGCCGAGGAGAGCAACGAGAC 60
DB 1 GACTAGTCTCTTGAAGTGTGAGAGAGAAAGCCGAGCCGAGGAGAGCAACGAGAC 60
QY 61 TGGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 61 TGGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 CCAACTGAGAGGTCCGAGATGAGAGAGCCGAGAGAGAGGAGGAGGAGGAGGAGGAG 180
DB 121 CCAACTGAGAGGTCCGAGATGAGAGAGCCGAGAGAGAGGAGGAGGAGGAGGAGGAG 180
QY 181 GGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 GGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 GGTCTGCTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 241 GGTCTGCTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 301 GAGCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 301 GAGCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 361 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 421 AGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

```

```

DB 421 AGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 481 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 541 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 541 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 601 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 661 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 721 GTTGTATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 721 GTTGTATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 781 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 GTGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 GTGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 CAGCTTCTCCGAGATTTCTGAGTACTCGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 901 CAGCTTCTCCGAGATTTCTGAGTACTCGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAG 960
QY 961 CCACTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 961 CCACTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 1021 GGTATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GGTATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GCACTCTCCGCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1081 GCACTCTCCGCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1141 GAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 GAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 GTGTAAGTCCCGGAGTGTGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 GTGTAAGTCCCGGAGTGTGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 CTGCTCTCTCTGAGTCTCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1261 CTGCTCTCTCTGAGTCTCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1321 GATCACTCAATTAACCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377
DB 1321 GATCACTCAATTAACCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377

```

RESULT 8

AAX24684

AAX24684 standard; cDNA; 1347 BP.

AAX24684;

21-JUN-1999 (first entry)

Human adipocyte-specific protein zs1g39 cDNA.

Db 1274 TCTCAGAGTCACTCAATAAAGACCTTCCTCAAAAAAAAAAAAAAAAAAAAA 1328

RESULT 10
AAFA4998
ID AAF44998 standard; cDNA; 1338 BP.
XX
AC AAF44998;
XX
XX 28-MAR-2001 (first entry)
XX
DE Human secreted protein related coding sequence SEQ ID NO: 103.
XX
XX Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
XX INTERCEPT 258; coronary disorder; olfactory disorder;
XX developmental disorder; pulmonary disorder; immunological disorder;
XX developmental disorder; kidney disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200078808-A1.
XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000MO-US16883.
XX
XX 18-JUN-1999; 99US-0336536.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Leiby KR, McKay C, Bossone S;
XX
XX WPI; 2001-050109/06.
XX
XX New nucleic acids for treating diseases and disorders, e.g.
XX atherosclerosis, infection, autoimmune diseases, obesity, ear
XX disorders, brain disorders, tumors, diabetes, arthritis, multiple
XX sclerosis and asthma -
XX
XX
XX Disclosure: Page 271; 332pp; English.
XX
XX The present invention provides the protein and coding sequences of the
XX human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
XX TANGO 281 and INTERCEPT 258. These are useful in the treatment of
XX coronary, pulmonary, olfactory, immunological, neurological,
XX developmental and kidney disorders.
XX
XX Sequence 1338 BP; 229 A; 422 C; 460 G; 227 T; 0 other;
XX

Query Match 94.0%; Score 1295; DB 22; Length 1338;
Best Local Similarity 99.5%; Pred. No. 2e-229;
Matches 1309; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 54 CCGAGACTGGGTGACCGGACGAGGAGGCGCTGCGCGGAGGAAAGCGGAGGCTGG 113
Db 15 CCGGACTGGGTGACCGGACGAGGAGGCGCTGCGCGGAGGAAAGCGGAGGCTGG 74

QY 114 AGACACCACTGAGAGGCTCCGAGTAGAGAGCGCCCGCAAGAGGCTCATCGGAGAGCC 173
Db 75 AGACACCACTGAGAGGCTCCGAGTAGAGAGCGCCCGCAAGAGGCTCATCGGAGAGCC 134

QY 174 GGGAGGGGGAGCTGCGAGAGAGACCCCGGCGCTCCGGGTCCCGGTGCGAGGCTATGAGGC 233
Db 135 GGGAGGGGGAGCTGCGAGAGAGACCCCGGCGCTCCGGGTCCCGGTGCGAGGCTATGAGGC 194

QY 234 CACTCTCTGCTGCTGCTCTCTGAGGCTGAGGCGCGGCTGCGCCCTCATGAGAGACA 293
Db 195 CACTCTCTGCTGCTGCTCTCTGAGGCTGAGGCGCGGCTGCGCCCTCATGAGAGACA 254

QY 294 AGATCCCGAGCTGCGCGGAGACCCCGGCTTTCAGAGGAGCGCGGAGCCAGCATGGCA 353
Db 255 AGATCCCGAGCTGCGCGGAGACCCCGGCTTTCAGAGGAGCGCGGAGCCAGCATGGCA 314

QY 354 GCCAGGCTTTCCTCGGAGCGCGGATGCGCGGAGCGGCGCGGAGCGCGGCGGCTCCGG 413
Db 315 GCCAGGCTTTCCTCGGAGCGCGGATGCGCGGAGCGGCGCGGAGCGCGGCGGCTCCGG 374

QY 414 GAGAGAAAGCGAGGCGGAGAGGCGGAGACTGCGGAGCCTCGAGGAGAGCCCGGAGCGG 473
Db 375 GAGAGAAAGCGAGGCGGAGAGGCGGAGACTGCGGAGCCTCGAGGAGAGCCCGGAGCGG 433

QY 474 GAGAGAGGCGGAGACCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 533
Db 434 GAGAGAGGCGGAGACCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 493

QY 534 GATCCGCTTCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 593
Db 494 GATCCGCTTCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 553

QY 594 TGCCCTTCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 653
Db 554 TGCCCTTCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 613

QY 654 TCACTGCGAGGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 713
Db 614 TCACTGCGAGGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 673

QY 714 GCTGCGAGTTGATCTGCTGAGAGAAATGCGAGATTCATGCTCTTCTTCAGTTTTCG 773
Db 674 GCTGCGAGTTGATCTGCTGAGAGAAATGCGAGATTCATGCTCTTCTTCAGTTTTCG 733

QY 774 GGGGATGAGCCCAAGCGAGCTGCTCTCGGAGGAGGAGCGGAGGAGGAGGAGGAGGAG 833
Db 734 GGGGATGAGCCCAAGCGAGCTGCTCTCGGAGGAGGAGCGGAGGAGGAGGAGGAGGAG 793

QY 834 ACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 893
Db 794 ACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 853

QY 894 CAGACAGACCTTCTCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 953
Db 854 CAGACAGACCTTCTCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 913

QY 954 CTTAGTGCACCTGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1013
Db 914 CTTAGTGCACCTGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 973

QY 1014 ACAACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1073
Db 974 ACAACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1033

QY 1074 GGGTGAAGCACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1133
Db 1034 GGGTGAAGCACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093

QY 1134 AGGTGTGAGCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1193
Db 1094 AGGTGTGAGCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1153

QY 1194 CTGGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1253
Db 1154 CTGGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1213

QY 1254 CTTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1313
Db 1214 CTTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273

QY 1314 TCTCAGAGTCACTCAATAAAGACCTTCCTCAAAAAAAAAAAAAAAAAAAAA 1368
Db 1274 TCTCAGAGTCACTCAATAAAGACCTTCCTCAAAAAAAAAAAAAAAAAAAAA 1328

RESULT 11
AAFA4999
ID AAF44999 standard; cDNA; 1338 BP.
XX

AA576911
ID AA576911 standard; cDNA; 1082 BP.
XX
AC AA576911;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12715.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX
DR P-PSDB; ABG12724.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 12715; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridization probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AA564197-AA594564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SO Sequence 1082 BP; 193 A; 334 C; 353 G; 202 T; 0 other;
Query Match 67.1%; Score 923.6; DB 23; Length 1082;
Best Local Similarity 96.5%; Pred. No. 4.4e-161;
Matches 1042; Conservative 0; Mismatches 24; Indels 14; Gaps 9;
QY 299 CCAGAGCTCTGCCCCGGGGGACCCCGCTTCCAGGCAAGCGGGCCACATGCGACGCG 358
DB 1 CCAGAGCTCTGCCCCGGGGGACCCCGCTTCCAGGCAAGCGGGCCACATGCGACGCG 60
QY 359 GGCTTGCAGGCGCGCGATGCGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 418
DB 61 GGCTTGCAGGCGCGCGATGCGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

QY 419 AAGGCGAGGCGCGGAGGCGCGGAGCTCCGCGGACCTTCAGGAGGACCCCGGCGCGAGGA 478
DB 121 AAGGCGAGGCGCGGAGGCGCGGAGCTCCGCGGACCTTCAGGAGGACCCCGGCGCGAGGA 180
QY 479 GAGCGGAGGACCCCGGCGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 538
DB 181 GAGCGGAGGACCCCGGCGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 240
QY 539 GCCTTCAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 598
DB 241 GCCTTCAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 300
QY 599 TTGACCGGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 658
DB 301 TTGACCGGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 360
QY 659 TGCCAGGCGGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 718
DB 361 TGCCAGGCGGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGG 420
QY 719 CAGTTGATCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778
DB 421 CAGTTGATCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 779 TGCCCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGAGG 838
DB 481 TGCCCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGAGG 540
QY 839 GTGTGGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 898
DB 541 GTGTGGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 600
QY 899 AGCAGCTTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 955
DB 601 AGCAGCTTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 956 TAGTGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1012
DB 661 TAGTGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 1013 GACAAACAGGATATCC-AGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1071
DB 721 GACAAACAGGATATCC-AGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 1072 T--GGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1127
DB 781 TTGGGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 1128 GCGATCAGGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1187
DB 841 GCGATCAGGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
QY 1188 GCTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1246
DB 901 GCTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 1247 CTCTCTCTCTGCT 1306
DB 961 CTCTCTCTCTGCT 1020
QY 1307 GCCCTTTTCTCA--GAGATCACTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1364
DB 1021 GCCCTTTTCTCAAGATCACTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080

Search completed: June 20, 2003, 06:55:51
Job time : 353 secs

Qy	4	CGCCAGGAGGCAACCAAGACCTGGGCTACCGGCAAGGAGGAGGCGCTGGCCGGGAGA	99
Db	11	CGAGAGGAGCCAAACCAAGACTGGGGTACCGCAAGGAGCGGCGCTGGCCGGGAGA	70
Qy	100	AGCGCGGGGGCTGGAGCACACCACTGAGGGGTCCGAGTAGGAGAGGCCCGAAGAG	15
Db	71	AGCGCGGGGGCTGGAGCACACCACTGAGGGGTCCGAGTAGGAGAGGCCCGAAGAG	13
Qy	160	GGCATTGGGAGACCCGGAGAGGGAGGACTCGAGAGGACCCCGGCGTCCGGGCTGCCG	215
Db	131	GGCATTGGGAGACCCGGAGAGGGAGGACTCGAGAGGACCCCGGCGTCCGGGCTGCCG	190
Qy	220	CAGCGCTATGAGGCACTCTTGCTCTGCTCTCTCTGAGGCTGGGCGGCCGACTGCCCC	279
Db	191	CAGCGCTATGAGGCACTCTTGCTCTGCTCTCTCTGAGGCTGGGCGGCCGACTGCCCC	25
Qy	280	ACTGAGCAGACAAAGATCCCAAGCTTGGCCGGGAGACCCCGGCGCTTCCAGGCAAGCC	339
Db	251	ACTGAGCAGACAAAGATCCCAAGCTTGGCCGGGAGACCCCGGCGCTTCCAGGCAAGCC	310
Qy	340	GGGCGCACTGGCAGCCAGGGCTTTCGGGCTCGAGTAGCCGCGCAAGCGCGC	399

```

Db      311 GGGCCACCATGCGAGCCAGGGCTTGCCTGGGCGCGCATGCGCCGCGACGCGCCGCGAGCGC 370
Qy      400 GCGCCGGGGGCTCCGGGAGAGAAAGCGAGGCGGGAGCGGGAGCTGCGGGAGCTTCGAGG 459
Db      371 GCGCGGGGCTCCGGGAGAGAAAGCGAGGCGGGAGCGGGAGCTGCGGGAGCTTCGAGG 430
Qy      460 GGAACCCCGGGCGCGGAGAGAGCGGGAGCTCCGCGGGCCGACCGGGCTTCGCGGGAGTG 519
Db      431 GGAACCCCGGGCGCGGAGAGAGCGGGAGCTCCGCGGGCCGACCGGGCTTCGCGGGAGTG 490
Qy      520 CTCGGTGCCTCCGCGAGCTCCGCTTCAGCGCGCAAGCGCTCCGAGAGCGGGGTCCTCCGCG 579
Db      491 CTCGGTGCCTCCGCGAGCTCCGCTTCAGCGCGCAAGCGCTCCGAGAGCGGGGTCCTCCGCG 550
Qy      580 GTCTGAGCGACCTTCCTCCCTTCGACCGCGTGTGTGAACGAGCAGGAGCATTTAGAGCGC 639
Db      551 GTCTGAGCGACCTTCCTCCCTTCGACCGCGTGTGTGAACGAGCAGGAGCATTTAGAGCGC 610
Qy      640 CGTCAACCGGCAAGTTCACTGCGAGGTGCTGGGGTCTACTTGTGCGCTGCATGCGCAC 699
Db      611 CGTCAACCGGCAAGTTCACTGCGAGGTGCTGGGGTCTACTTGTGCGCTGCATGCGCAC 670
Qy      700 CGTCAACCGGCGCAGCTGCTGAGTTGATCTGGTGAAGATGCGAATCCATTGCGCTCTT 759
Db      671 CGTCAACCGGCGCAGCTGCTGAGTTGATCTGGTGAAGATGCGAATCCATTGCGCTCTT 730
Qy      760 CTTCAGATTTCCTGGGGGGTGGCGCCAGCGCTTCGCTCTCGGGGGGGGCGCATGGTAG 819
Db      731 CTTCAGATTTCCTGGGGGGTGGCGCCAGCGCTTCGCTCTCGGGGGGGGCGCATGGTAG 790
Qy      820 GCTGAGAGCTGAGAGCAAGTGTGGGTGAGAGGTGGGTGGGTGATCAATTGGGATCTA 879
Db      791 GCTGAGAGCTGAGAGCAAGTGTGGGTGAGAGGTGGGTGGGTGATCAATTGGGATCTA 850
Qy      880 TGCCAGCATCAAGACAGACAGACCTTCTCCGATTTCTGTGTATCTCCGACTGGCAGAG 939
Db      851 TGCCAGCATCAAGACAGACAGACCTTCTCCGATTTCTGTGTATCTCCGACTGGCAGAG 910
Qy      940 CTCCCGAGCTTCCTGATGATGCGCACTGCAAGTGAAGCTCATGCTCTCACTCTAGAAAG 999
Db      911 CTCCCGAGCTTCCTGATGATGCGCACTGCAAGTGAAGCTCATGCTCTCACTCTAGAAAG 970
Qy      1000 AGGCTGTGAGGCTGACCAACAGGTCATCCAGAGAGGCTGGCGCCCTGGATATATGTGA 1059
Db      971 AGGCTGTGAGGCTGACCAACAGGTCATCCAGAGAGGCTGGCGCCCTGGATATATGTGA 1030
Qy      1060 TGACTAGAGAGGTGGGGTGAAGCACTTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1119
Db      1031 TGACTAGAGAGGTGGGGTGAAGCACTTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1090
Qy      1120 GAGCTGTGCGAGTCAAGGTCTGGCAGAGATGGGCAAGTGTGATTTCTGCGCCCAACAGAG 1179
Db      1091 GAGCTGTGCGAGTCAAGGTCTGGCAGAGATGGGCAAGTGTGATTTCTGCGCCCAACAGAG 1150
Qy      1180 AGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1239
Db      1151 AGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1210
Qy      1240 TGGGGTGTCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1299
Db      1211 TGGGGTGTCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1270
Qy      1300 GGGGGCGGGCCCTTTCTCAGAGATCACTCAATTAACCTTAAGAACCTCATTAATAAAAAA 1359
Db      1271 GGGGGCGGGCCCTTTCTCAGAGATCACTCAATTAACCTTAAGAACCTCATTAATAAAAAA 1330
Qy      1360 AAAAAA 1366
Db      1331 AAAAAA 1337

```

RESULT 2
US-09-336-536-1

```

; Sequence 1, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-336-536-1

Query Match      94.2%; Score 1296.6; DB 4; Length 1338;
Best Local Similarity 99.6%; Pred. No. 7,76-252;
Matches 1310; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy      54 CCAGGACTGGGGTGAACGCGCAGGGCAAGGGGCGCTTGGCGGGGAGAAAGCGGGGCTGG 113
Db      15 CCGGACTGGGGTGAACGCGCAGGGCAAGGGGCGCTTGGCGGGGAGAAAGCGGGGCTGG 74
Qy      114 AGCACCACTGAGAGGTCCGAGTGAAGAGCGCCCGAAGAGGCAATCCGGGGAGCC 173
Db      75 AGCACCACTGAGAGGTCCGAGTGAAGAGCGCCCGAAGAGGCAATCCGGGGAGCC 134
Qy      174 GGGAGGGGGAGCTGCGAGAGAGACCCCGGCGTCCGGGCTCCGGTGAAGCGCTATGAGGC 233
Db      135 GGGAGGGGGAGCTGCGAGAGAGACCCCGGCGTCCGGGCTCCGGTGAAGCGCTATGAGGC 194
Qy      234 CACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 293
Db      195 CACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 254
Qy      294 AGATCCCGAGCTTGTGCGCGGGGCAAGCCCGGCTTCAGAGCAAGCGGGGCAATGGCA 353
Db      255 AGATCCCGAGCTTGTGCGCGGGGCAAGCCCGGCTTCAGAGCAAGCGGGGCAATGGCA 314
Qy      354 GCCAGGCTTGTGCGCGGGGCGAGATGCGCGGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 413
Db      315 GCCAGGCTTGTGCGCGGGGCGAGATGCGCGGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 374
Qy      414 GAGAGAAAGGCGAGGGCGGGAGAGCGGGGCACTGCGGGGCACTGAGGGGAAACCGGGGCGG 473
Db      375 GAGAGAAAGGCGAGGGCGGGAGAGCGGGGCACTGCGGGGCACTGAGGGGAAACCGGGGCGG 433
Qy      474 GAGAGAGCGGGAGCCCGCGGGGCAAGCGGGGCTGCGGGGAGTCTCGGTGCTTCGCG 533
Db      434 GAGAGAGCGGGAGCCCGCGGGGCAAGCGGGGCTGCGGGGAGTCTCGGTGCTTCGCG 493
Qy      534 GATCCGCTTCAAGGCGCAAGCGCTCGAGAGCGGGGCTCGCGCGCTGAGAGCAAGCT 593
Db      494 GATCCGCTTCAAGGCGCAAGCGCTCGAGAGCGGGGCTCGCGCGCTGAGAGCAAGCT 553
Qy      594 TGCCCTTCAAGCGGGTCTGTGAACGAGGAGCAATTAAGAGCGCGTCAAGCGGCAAGT 653
Db      554 TGCCCTTCAAGCGGGTCTGTGAACGAGGAGCAATTAAGAGCGCGTCAAGCGGCAAGT 613
Qy      654 TCACTGCGAGGTCCTGGGGTCTACTACTTTCGCGCTCAATGCGACCTTCAAGCGGCA 713
Db      614 TCACTGCGAGGTCCTGGGGTCTACTACTTTCGCGCTCAATGCGACCTTCAAGCGGCA 673
Qy      714 GCTGCAATTGATCTGTGAAGATGCGAATTCATATGCTCTTCTTCAGATTTTGG 773
Db      674 GCTGCAATTGATCTGTGAAGATGCGAATTCATATGCTCTTCTTCAGATTTTGG 733
Qy      774 GGGGGTGGCCCAAGCGCTCGCTCTCGGGGGGGGCGATGTGAAGCTGAAGCTGAAGG 833
Db      734 GGGGGTGGCCCAAGCGCTCGCTCTCGGGGGGGGCGATGTGAAGCTGAAGCTGAAGG 793

```


QY	834	ACCAAGTGTGGGTGCAGGTGGGGTGGGAGCTAACTATTGGATCTATGCAACATCAGA	893
Db	794	ACCAAGTGTGGGTGCAGGTGGGTGGGAGCTAACTATTGGATCTATGCAACATCAGA	853
QY	894	CAGACAGCACTTTCTCCGAGATTTCTGTGTACTCCGACATGAGACAGCTCCCGACTTTTG	953
Db	854	CAGACAGCACTTTCTCCGAGATTTCTGTGTACTCCGACATGAGACAGCTCCCGACTTTTG	913
QY	954	CTTAGTGCCACATCGCAAAATGATGAGCTCATCTCTCACTCTCTAAGAGAGAGGTGTGAGGCTG	1013
Db	914	CTTAGTGCCACATCGCAAAATGATGAGCTCATCTCTCACTCTCTAAGAGAGAGGTGTGAGGCTG	973
QY	1014	ACAACCAAGTTCATCCAGAGAGGCTGAGCCCCCTGTGAATATTTGTGAATGACTAGAGGAGTG	1073
Db	974	ACAACCTGTTCATCCAGAGAGGCTGAGCCCCCTGTGAATATTTGTGAATGACTAGAGGAGTG	1033
QY	1074	GGGTAGAGACCTCTCCGTCTGTCTGTGTGGCAAGAAATGGAAACATGTGCTGTCTCCGATC	1133
Db	1034	GGGTAGAGACCTCTCCGTCTGTCTGTGTGGCAAGAAATGGAAACATGTGCTGTCTCCGATC	1093
QY	1134	AGGTCTGTGCAGCATGGGGCAGTGGCTGTGAATTTCTGCCAAGACAGAGAGATGTGCTGTG	1193
Db	1094	AGGTCTGTGCAGCATGGGGCAGTGGCTGTGAATTTCTGCCAAGACAGAGAGATGTGCTGTG	1153
QY	1194	CTGGCAAGTGAATGCCCCCAATGTCTCTGTGTGCAGAGAGCCACAGGTGTGGGTCTCTTT	1253
Db	1154	CTGGCAAGTGAATGCCCCCAATGTCTCTGTGTGCAGAGAGCCACAGGTGTGGGTCTCTTT	1213
QY	1254	CTGTGTCTCTGTCTTCTGTGAATCTGCCACACCCCTCTGTCTCTGTGGGGCAGGAGCTTTT	1313
Db	1214	CTGTGTCTCTGTCTTCTGTGAATCTGCCACACCCCTCTGTCTCTGTGGGGCAGGAGCTTTT	1273
QY	1314	TCTGAGAGATCACTCAATTAACCTTAAGAAACCTCATTAATAAAAAAAAAAAAAAAAAA	1368
Db	1274	TCTGAGAGATCACTCAATTAACCTTAAGAAACCTCATTAATAAAAAAAAAAAAAAAAAA	1328

RESULT 3
 US-09-336-536-2
 : Sequence 2, Application US/09336536
 : Patent No. 6406884
 : GENERAL INFORMATION:
 : APPLICANT: Leiby, K.
 : APPLICANT: McKay, C.
 : APPLICANT: Bossone, S.
 : TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 : FILE REFERENCE: 7853-144
 : CURRENT APPLICATION NUMBER: US/09/336,536
 : CURRENT FILING DATE: 1999-06-18
 : NUMBER OF SEQ ID NOS: 75
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 2
 : LENGTH: 728
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-336-536-2

Query Match	52.1%;	Score 717;	DB 4;	Length 728;
Best Local Similarity	99.9%;	Pred. No. 1.4e-135;		
Matches 728; Conservative	0;	Mismatches 0;	Indels 1;	Gaps 1

OY	227	ATGAGGCACTCTCGTCTGCGCGCTCTGAGCGCTGAGCGAGCGCGCTCGGCCCATCTGAGC	286
Db	1	ATGAGGCACTCTCGTCTGCGCTCTGAGCGCTGAGCGAGCGCGCTCGGCCCATCTGAGC	60
OY	287	GACCAACAGATCCCAAGCTCTGCGCGGAGCAACCCGCGCTTCACAGCAACCGCGGCAC	346
Db	61	GACCAACAGATCCCAAGCTCTGCGCGGAGCAACCCGCGCTTCACAGCAACCGCGGCAC	120
OY	347	CATGCGAGCAGGCGCTTGCCGAGCGCGAGTAGCGCGAGCGCGCCGACGCGCGCGCG	406
Db	121	CATGCGAGCAGGCGCTTGCCGAGCGCGAGTAGCGCGAGCGCGCGCCGACGCGCGCGCG	180

QY	407	GCTCCGGAGAGAAAAGCGAGGCGGGAGGCGGGACCTCCGGACCTTCGAGGGACCCC	466
Db	181	GCTCCGGAGAGAAAAGCGAGGCGGGAGG- GGGACTTCGGGACCTTCGAGGGACCCC	239
QY	467	GGGCGCGAGAGAGGCGGGACCCCGGGGGCCACCGGGGCTGGCGGGGAGTGTCTGGTG	526
Db	240	GGGCGCGAGAGAGGCGGGACCCCGGGGGCCACCGGGGCTGGCGGGGAGTGTCTGGTG	299
QY	527	CTCCGCGATCCGCTTCAGCGCCAAAGCGCTCCGAGAGCGGGATGCTCCGCGCTGCAC	586
Db	300	CTCCGCGATCCGCTTCAGCGCCAAAGCGCTCCGAGAGCGGGATGCTCCGCGCTGCAC	359
QY	587	GCAACCTTGCCCTTGCACCGCGTGTCTGTGAACGAGCAGGAGACATTACGACGCGTACC	646
Db	360	GCACCTTGCCCTTGCACCGCGTGTCTGTGAACGAGCAGGAGACATTACGACGCGTACC	419
QY	647	GGCAGATTCACTGCGAGGTGCTGAGGGGTCTACTTGGCGGTCCGATGCACGCGTAC	706
Db	420	GGCAGATTCACTGCGAGGTGCTGAGGGGTCTACTTGGCGGTCCGATGCACGCGTAC	479
QY	707	CGGCGCAGCTGCAAGTTGATCTGTGTGAAGATGCGGATTCATTGCTCTTCTTCCAG	766
Db	480	CGGCGCAGCTGCAAGTTGATCTGTGTGAAGATGCGGATTCATTGCTCTTCTTCCAG	539
QY	767	TTTTTCGGGGGGTGGCCCAAGCTTCGCTCTCGGGGGGGCCATGTTGAGGCTGAG	826
Db	540	TTTTTCGGGGGGTGGCCCAAGCTTCGCTCTCGGGGGGGCCATGTTGAGGCTGAG	599
QY	827	CCTGAGGACCAAGTGTGGGTGCAAGTGTGGGTGTGGGTGACTACATTGGGATCTATGCGAG	886
Db	600	CCTGAGGACCAAGTGTGGGTGCAAGTGTGGGTGTGGGTGACTACATTGGGATCTATGCGAG	659
QY	887	ATCAAGACAGACAGCACTTCTCGGATTTCTGTGTACTTCGACTGGCAGAGCTTCCCA	946
Db	660	ATCAAGACAGACAGCACTTCTCGGATTTCTGTGTACTTCGACTGGCAGAGCTTCCCA	719
QY	947	GTCCTTGCT 955	
Db	720	GTCCTTGCT 728	

```

RESULT 4
US-09-336-536-8
: Sequence 8, Application US/09336536
: Patent No. 6406884
: GENERAL INFORMATION:
: APPLICANT: Leiby, K.
: APPLICANT: McKay, C.
: APPLICANT: Bossone, S.
: TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
: FILE REFERENCE: 7853-144
: CURRENT APPLICATION NUMBER: US/09/336,536
: CURRENT FILING DATE: 1999-06-18
: NUMBER OF SEQ ID NOS: 75
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 8
: LENGTH: 1263
: TYPE: DNA
: ORGANISM: Mus musculus
US-09-336-536-8

```

	Query Match	Similarity	Score	692.8	DB	4	Length	1253
	Best Local	Similarity	78.0%	Pred.	No. 1-1e-150			
	Matches	917	Conservative	0	Mismatches	23	Indels	22
								Gaps . 6
OY	189	GAGAGAACCCCGCGGCTCGGCTCCCGGTGCGACGGCTTATGAGGCCACTCTCGTCTGC	248					
Db	97	GAGAGGTTCTCGAGAGTCTGAGCCCTTCGGGGGTCAACCACTATAGGCCCACTTTCGGCCCTTC	156					
OY	249	TGCTCTTGAGGCTTGAGCGCGCTCGGCTCGCCCCACTCGACGACGACGACGCTCT	308					
Db	157	TGCTTCTGAGGCTTGAGTGTGACAGGCTCTCTCTCTCGACGACGACGACGACGCTCT	216					

Oy	309	GC	CCGCGGGGCA	CC	CGGCTTCT	CA	GGCAGC	CG	CGGGCC	CA	CA	ATG	GA	GC	CA	GGG	CTT	GC	CG	368
Db	217	GT	CCCGGCA	CC	CGGCTTCT	CA	GGCAGC	CG	CGGGCC	CA	CA	ATG	GA	GC	CA	GGG	CTT	GC	CG	276
Oy	369	GC	CCGCGATG	CG	CGCGA	CGG	CCGCA	CGG	CGCG	CG	CG	CGG	CG	CG	CG	CG	CG	CG	CG	428
Db	277	GC	CCGTGAC	CG	CGCGATG	CGG	CCGCA	CGG	CGCGA	CG	CG	CGG	CG	CG	CG	CG	CG	CG	CG	336
Oy	429	GG	GGGAGG	CG	CGGAC	CTG	CGAGG	GGG	AC	CC	CGG	CG	CG	CG	CG	CG	CG	CG	CG	488
Db	337	GG	GGGAGG	CG	CGGAC	CTG	CGAGG	GGG	AC	CC	CGG	CG	CG	CG	CG	CG	CG	CG	CG	396
Oy	489	CC	GGCGGG	CC	CGGCG	CTG	CGG	GGG	AGT	CT	CG	AGT	GC	CT	CG	CG	CG	CG	CG	548
Db	397	CC	ATG	GGGGG	CG	CTAT	GGG	CG	CGG	GGG	AGT	CT	CG	AGT	GC	CT	CG	CG	CG	456
Oy	549	CC	AGCG	CT	CG	AG	AG	CG	CGG	TG	CT	CG	CG	CG	CG	CG	CG	CG	CG	608
Db	457	CC	AGCG	CT	CG	AG	AG	CG	CGG	TG	CT	CG	CG	CG	CG	CG	CG	CG	CG	516
Oy	609	TG	CTG	GT	GA	CA	GA	GGG	GA	CA	TT	TA	CG	AG	CG	CG	CG	CG	CG	668
Db	517	TG	CTG	GT	GA	CA	GA	GGG	GA	CA	TT	TA	CG	AG	CG	CG	CG	CG	CG	576
Oy	669	CT	GGG	GT	CT	AC	TT	CG	CG	CG	TC	AG	CG	CG	CG	CG	CG	CG	CG	728
Db	577	CT	GGG	GT	CT	AC	TT	CG	CG	CG	TC	AG	CG	CG	CG	CG	CG	CG	CG	636
Oy	729	TG	GT	GA	GA	AT	AG	CG	GA	AT	TC	AT	TG	CT	TT	CT	CG	AG	TT	788
Db	637	TG	GT	GA	GA	AT	AG	CG	GA	AT	TC	AT	TG	CT	TT	CT	CG	AG	TT	696
Oy	789	C	AG	CT	CG	CT	CG	CG	GG	GG	GG	CG	CA	AT	AG	GA	GG	CT	GA	848
Db	697	C	AG	CT	CG	CT	CG	CG	GG	GG	GG	CG	CA	AT	AG	GA	GG	CT	GA	756
Oy	849	AG	GT	GG	GT	GG	GT	GA	CT	AC	AT	TG	GC	AT	TC	AT	GC	AG	CA	908
Db	757	AG	GT	GG	GT	GG	GT	GA	CT	AC	AT	TG	GC	AT	TC	AT	GC	AG	CA	816
Oy	909	C	CG	GA	TT	CT	GT	GT	GA	CT	CG	CA	CT	CG	CG	CG	CG	CG	CG	968
Db	817	C	CG	GA	TT	CT	GT	GT	GA	CT	CG	CA	CT	CG	CG	CG	CG	CG	CG	876
Oy	969	A	AG	GT	AG	CT	AT	GT	CT	CA	CT	CG	TA	GA	AG	AG	GG	GT	GA	1028
Db	877	A	AG	GT	AG	CT	AT	GT	CT	CA	CT	CG	TA	GA	AG	AG	GG	GT	GA	936
Oy	1029	A	GA	AG	GG	CT	GG	CC	CC	CT	CG	TA	AT	TG	GA	AT	GA	TA	GA	1088
Db	937	A	GA	AG	GG	CT	GG	CC	CC	CT	CG	TA	AT	TG	GA	AT	GA	TA	GA	996
Oy	1089	C	GT	CT	GT	GT	CG	GA	GA	AT	GG	GA	CA	GT	GG	CG	GT	CT	GA	1148
Db	996	A	GT	CC	CA	CT	GT	CG	GA	GA	AT	GG	GA	CA	GT	GG	CG	GT	CT	1049
Oy	1149	G	GG	CA	GT	GG	AT	TT	CT	GG	CC	AG	CA	GA	AG	AG	GT	CT	GT	1208
Db	1050	G	GG	CA	GT	GG	AT	TT	CT	GG	CC									

```

RESULT 5
US-09-188-930-218
; Sequence 218, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muirson, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 218
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-218

Query Match          44.0%; Score 606; DB 3; Length 1001;
Best Local Similarity 78.6%; Pred. No. 2,78-113;
Matches 739; Conservative 0; Mismatches 195; Indels 6; Gaps 1;

QY      66  TGACGGCAGGCGCAGGGGGCGCTGGCGCGGAGAAAGCCGGGGGCTGAGCACAACCAAC 125
DB      31  TGTACGAGAGGGGCGAGGGGCTGCTGGCTGGTGGGGTGAAGTGGAGACAGGGGCGACGAGAGG 90

QY      126  TGGAGGGGTCCGGAGTACCGGACCGCCCGAAGAGAGCCATCGGGAGCGCGGAGGGGGAGC 185
DB      91  GTCTGAGGAAGCCATTCMAAAGCCAGCAGCTGGAGAGACTGGGAGGCGGGAAAGGGCTTAC 150

QY      186  TG-----CGAAGAGAACCCCGGCGCTCCGGGCTCCCGGTGCGAGGCTATGAGGCCACTCC 239
DB      151  AGACTTACAAAGAGAGATCTTGCGCTCTGGGCTCTCTGGGTATACCATAGAGGCCACTTC 210

QY      240  TCGTCTGCTGCTCTCTGGGCTTGCGGGCCGAGCTCGCCCACTGAGCAGCAACAAGATCC 299
DB      211  TTGGCCCTGCTGCTCTGGGCTTGCGCATCAGGCTCTCTCTCTGAGCAGCAACAAGATCC 270

QY      300  CCAAGCTCTGCGCGGGGAGCCCGGCGCTTCGAGGAGCAGCGGGGCGACCAATGGACAGAG 359
DB      271  CCAAGCTGTGTCTCCGGGCGAGCCGGGCTTCCAGGAGCAACAAGGCAACAAGCGGACCAAG 330

QY      360  GCTTGGCGGGCCGCGATGAGCGCGCAGCGCGCAGCGCGGCGCTCCGGGAGAGA 419
DB      331  GCTTGGCGGGCGGAGAGCGGCGGTGATGGCGGCGAGCGGTGACGCCGAGGCTCCGGGAGAGA 390

QY      420  AAGCGAGGGCGGGAGGCGCGGAGCTCGCGGAGACTCGAGGGAGACCTCGGGGCTCGGAGAG 479
DB      391  AAGCGAGGGCGGGAGACCGGAGCTACTCGGCGCAGCTGGGAGACCGGGGCGGGTGGAG 450

QY      480  AGGCGGAGCCCGGAGGCGCAACCGAGGCTGCGGGGAGGTGCTCGAGTGCCTCGGAGATCG 539
DB      451  AAGCGAGACTCTGTGGGGGCTTATCGGAGCTCGGGGAGAGTGTCTCGATGCCCCACGATGAG 510

QY      540  CTTTCAGCGCGCAAGCGCTTCGAGAGACCGAGGTGCTTCGCGCGTCTGAGCGCACCTTGGCCT 599
DB      511  CTTTCAGTGCAGACGATCAGAGAGGCGGGATCTTCGCGCGAGCGAGCACACCTTACCTT 570

QY      600  TCGACCGCGTCTGTGTGAACGAGAGGAGCACTTACGAGCCGCTACCGGGAGATTACCT 659
DB      571  TCGACCGGTGTCTCAATGAGAGAGGAGCACTTACGATGCCATACCGGCAAGTTCACT 630

QY      660  GCCAGGTGCTCGGGGTCTACTACTTCCCGGTCATGCGCACCGTCTAACCGGGCGACGCTGC 719
DB      631  GCCAAGTGCCTGGTGTCTACTACTTTCGTGCATGCGCACGCTCTAACCGGGCGACGCTAC 690

QY      720  AGTTTATCTGTGATGAAGATGGCGAATTCATGCTCTTTTTCGAGTTTTCGGGGGGT 779

```


NAME/KEY: modified base
LOCATION: all "n" positions
OTHER INFORMATION: n=a, c, g, or t
US-09-336-536-74

Query Match 18.7%; Score 257.8; DB 4; Length 601;
Best Local Similarity 72.0%; Pred. No. 1.8e-43;
Matches 365; Conservative 0; Mismatches 134; Indels 8; Gaps 2;

```
QY 66 TCACGGAGGCGAGGGGGGCGCTGGCCGGGAGAGACCGGGGGCTGAGACACACAC 125
DB 196 TGTCAAGGAGGAGGAGGGGCTGCTGTGGGTGAGAGAGGAGGAGGAGGAGGAG 155
QY 126 TGAAGGGTCCGAGTAGAGAGCGCCCGAGAGAGGCATTCGGGGAGCCGGGAGGGGAGC 185
DB 156 GTCGTAGGAAGCATTCAAAGGAGAGAGAGCTGGGAGAGAGCTGGGAGAGCGGGAGGCTTAC 215
QY 186 TG-----CGAGAGACCCCGGGGCTCCGGGCTCCCGGTCACGCTATAGGCCACTCC 239
DB 216 AGACTACAGAGAGAGATCTGGGCGCTCCGGGCTCCCGGTCATCAACATAGAGGCCACTTC 275
QY 240 TCGTCTGCTGCTCTGAGGCTGGAGCGGGCGGCTGGCCCGGCTGAGAGAGAGAGATCC 299
DB 276 TGGCTCTGCTGCTCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 335
QY 300 CCAGCTCTGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
DB 336 CCAGCTCTGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 395
QY 360 GCTTGGCGGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
DB 396 GCTTGGCTGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 455
QY 420 AAGCGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
DB 456 AAGCGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 515
QY 480 AGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
DB 516 AGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
QY 540 CCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 566
DB 574 GCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
```

RESULT 10

US-09-188-930-26
Sequence 26, Application US/09188930A

GENERAL INFORMATION:
PATENT NO. 6150502
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Orniet, Rene
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq For Windows Version 3.0
SEQ ID NO 26
LENGTH: 393
TYPE: DNA
ORGANISM: Rat
US-09-188-930-26

Query Match 10.9%; Score 149.6; DB 3; Length 393;
Best Local Similarity 69.1%; Pred. No. 8.6e-22;
Matches 250; Conservative 0; Mismatches 104; Indels 8; Gaps 3;

```
QY 66 TCACGGAGGCGAGGGGGGCGCTGGCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 125
DB 33 TGTCAAGGAGGAGGAGGGGCTGCTGTGGGTGAGAGAGGAGAGAGAGAGAGAGAGAGAG 92
QY 126 TGAAGGGTCCGAGTAGAGAGCGCCCGAGAGAGGCATTCGGGGAGCCGGGAGGGGAGC 185
DB 93 GTCGTAGGAAGCATTCAAAGGAGAGAGAGCTGGGAGAGAGCTGGGAGAGCGGGAGGCTTAC 152
QY 186 TG-----CGAGAGACCCCGGGGCTCCGGGCTCCCGGTCACGCTATAGGCCACTCC 239
DB 153 AGACTACAGAGAGAGATCTGGGCGCTCCGGGCTCCCGGTCATCAACATAGAGGCCACTTC 212
QY 240 TCGTCTGCTGCTCTGAGGCTGGAGCGGGCGGCTGGCCCGGCTGAGAGAGAGAGATCC 299
DB 213 TGGCTCTGCTGCTCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 272
QY 300 CCAGCTCTGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
DB 273 CCAGCTCTGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 332
QY 360 GCTTGGCGGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
DB 333 GCTTGGCTGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 390
```

RESULT 11

US-08-463-911-1
Sequence 1, Application US/08463911

PATENT NO. 5869330

GENERAL INFORMATION:

APPLICANT: Scherer, Philipp E.

APPLICANT: Lodish, Harvey F.

TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED

TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,911

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: WHI95-05

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1276 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 46..786

US-08-463-911-1

Query Match	6.6%;	Score 90.6;	DB 2;	Length 1276;
Best Local Similarity	51.5%;	Pred. No. 7.3e-10;		
Matches 316;	Conservative 0;	Mismatches 279;	Indels 18;	Gaps 4.

Qy	333	GCACGCCGGGCGCACCATGGCAGACCGGGCTTGGCCGGGCGCGCATGGCCGCGACGGCCGGG	392
Db	170	GGATGGCAGGCATCCACGACCACTCTGGCGCACAATGGCACAACAGGCCCGTGAATGGCAGAG	229
Qy	393	ACGGCGCGCCCTCGGGGCTCTCCGGGAGAAAGCGCGAGGGCGGGAGCGCGGACCTGCCGGGAC	452
Db	230	ATGGCACCTCCCTGGAGAGAAAGGAGAAAGAGAGATGCAGGTCTTCTTGATCTTAAAGGGTG	289
Qy	453	CTGCAGGGGGACCCCGGGCCGCGAGAGAGGGCGGACCCCGGGGGCCACCGGGGCTCGCCG	512
Db	290	AGACAGAGAGATGTTGGAAATGACAGAGGCTGAAAGGGCCACGGGGCTTCCCGGAAACCCCTG	349
Qy	513	GGAGATGTCGGATGCTCCCGGATCCCGGATCCGCTTACG----CGCCAGCGCTCCGAGAGCCGG	568
Db	350	GCAGAGAAAGAGAGAGCTGAGAAAGCCGCTTATATGATGCTCAGGGTTCAAGTGGGGC	409
Qy	569	GTCGCTCCGCGCTCG-----ACGACCCCTTGCCCTTCGACCGCGCTGCTGGAGACG	620
Db	410	TGGAGACCCCGCGCACTGTTCCTCCAAATGTAACCATTCGCTTAAAGATCTTCTTACMAACC	469
Qy	621	AGCAGGACATTACGACGCGCTCACCGGCAATTCACCTCGCCAGAGGCTCGGGGATCTACT	680
Db	470	AACGAAATCATTTATACGCGACGACACTGGGCAATTTCTACTGCAACATTCGGGAACTTACT	529
Qy	681	ACTTCGCGGTCATGCAACCGATCTAACCGGGCAGGCTGCAATTGATCTGGTGAAGATG	740
Db	530	ACTTCTCTTACCAATCAACGAGGTGATCAATGAAGAATGGAAGGTGAGGCTCTTCAAGAAAG	589
Qy	741	GCGAATCATATGACCTCTTCTTCCAGTTTTCGGGGGGGTGGGCCCAAGCCAGCCTGCTCT	800
Db	590	ACAAGGCGGTCTCTTTCACCTACGACCAAGATACGAAAG--AAATGGAGCCAGGCTT	646
Qy	801	CGGGGGGGGCGCATGGTGAAGGCTGAGGCTGAGGACCAATGTGGTGCAGATG--GGTG	857
Db	647	CTGGCTCTGTGTCTCTTCATCTGGAGGTGGAGACCAATGTGGGCTCCAGGTGTATGGGG	706
Qy	858	TGGGTGACTACATTGGCATCTATGCGCAGATCAAGACAGACAGCACTTCTCGGATTTTC	917
Db	707	ATGGGAGACCAATGACTCTATGCAAGTAAAGTCACAGCACTTACATTAACTGCGCTTTC	766
Qy	918	TGGGTACTCCGA	930
Db	767	TTCTTACCATGA	779

RESULT 12 12-911-6
 US-08-463-911-6
 Sequence 6, Application US/08463911
 Patent No. 5863330
 GENERAL INFORMATION:
 APPLICANT: Scherer, Philipp E.
 APPLICANT: Lodish, Harvey F.
 TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
 TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: Two Militia Drive
 CITY: Lexington
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,911

```

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH195-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEO ID NO.: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1113 base pairs
TYPE: nucleic acid
STANDARDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 73..804
US-08-463-911-6

```

Query Match	6.0%;	Score	82.2;	DB	2;	Length	1313;
Best Local Similarity	50.1%;	Pred.	0;	3.6e-08;			
Matches	316;	Conservative	0;	Mismatches	303;	Indels	12;
						Gaps	4;

Oy	322	CCGGCCCTTCAGAGGCAGCGCCGGGCCACCATGCAATGCAAGGCTTTGCCGGGCGCGGAATGACC	380
Db	194	CGGACATCCCAAGGGCATCCGGGCCCATTAATATGAGGGGCCCAAGGCCGTGATGCGAGAGATGGCA	253
Oy	381	GGCAGCGCCCGCAGCGCGCGCCCGGGCTCCGGGAGAGAAAGCGAGGAGCGGAGGCCG	440
Db	254	CCCTTGATGAGAGGGGTGAGAAAGAGATCCAGGCTCTTAATGGCTCTAAGAGAGACATCG	313
Oy	441	GACTGCCGGGACCTTGAGAGGACCCCGGGCCGCGAGAGAGAGGGGGACCCGCGGGGCCCA	500
Db	314	GTGAACCGGATATCCGGGCGTGAAGGTCTCCCAAGGCTTTTCGGGGAATCCAAAGCAGGA	373
Oy	501	CCGGCGCTCCCGGGAGATGCTCGGTCCTCCGCGATCCGCTTCAGCGCCAGCGCTCCG	560
Db	374	AAGGAAACCTTGAGAGAGGTGCTATGATATACCGCTCAGACTTCAGTGT---GGGATTGG	430
Oy	561	AGAGCCGGGTGCTTCGCGCGTCTGACGCACTTTGCCCTTCGACCGGTGCTGTGAACG	620
Db	431	AGACTTACGTTACTATCC---CCAACATGGCCATTCGCTTTACCAAGATCTTCAACATC	487
Oy	621	AGAGAGGAATTAACAGCCGCTCACCGGCAAGTTCACTCGCAGGTGCTCGGGGTCACT	680
Db	488	AGCAAAACCACTATATATGCTCAGTGGTAATTCACATGCAATTCCTGGGCTGTACT	547
Oy	681	ACTTCCGCTGCATGACCAACGCTTACCCGGGCGAGCTCGAGTTGATCTGTGAAGATG	740
Db	548	ACTTTGCTACACATCAACAGCTATATATGAAGATGTAAGGTCAAGCTCTTCAAGAAAG	607
Oy	741	GCGAATCCATTGCTCTTTCTTCCAGTTTTCGGGGGGTGGCCCAAGCAGCTCGCTCT	800
Db	608	ACAAAGCTATGCTCTTCACTATATGATACGACACAGGAAAT--AAATGTGACCAAGGCTT	664
Oy	801	CGGAGGGGGGCATATGTAAGGCTGAGAGCTGAGAGCAAGATGTGGGTGCAAGTGGATG-	859
Db	665	CGGACTCTGTGCTCTGCACTTGGAGAGTGGGCGACCAAGTCTGTGCTCAAGTGTATGGGG	724
Oy	860	--GGTACTAATTTGGCATCTATGCGACAGATCAAGACAGACAGACACTTTCGGAATTC	917
Db	725	AAGAGAGCGTATGGAATCTATATGCTGATATATGACATGATCACTCACTTCAAGGCTTTC	784
Oy	918	TGATGTATCCCATGGCAGACAGCTCCAGT	948
Db	785	TTCTCTACATGACCAACATGATACCACT	815

RESULT 13
US-09-140-804-9

RESULT 14 .
US-09-118-408-23
; Sequence 23, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 08:11:50 ; Search time 224 Seconds
(without alignments)
9020.747 Million cell updates/sec

Title: US-09-944-403-41
Perfect score: 1377
Sequence: 1 gactagctctcttgagctc.....aaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1377	100.0	1377	9	US-09-944-413-41 Sequence 41, App1
2	1377	100.0	1377	9	US-09-944-403-41 Sequence 41, App1
3	1377	100.0	1377	9	US-09-944-896-41 Sequence 41, App1
4	1377	100.0	1377	9	US-09-944-944-41 Sequence 41, App1
5	1377	100.0	1377	9	US-09-944-907-41 Sequence 41, App1
6	1377	100.0	1377	9	US-09-944-929-41 Sequence 41, App1
7	1377	100.0	1377	9	US-10-028-072-361 Sequence 361, App
8	1377	100.0	1377	9	US-10-121-049-361 Sequence 361, App
9	1377	100.0	1377	9	US-10-123-904-361 Sequence 361, App
10	1377	100.0	1377	9	US-10-140-470-361 Sequence 361, App
11	1377	100.0	1377	9	US-10-175-746-361 Sequence 361, App
12	1377	100.0	1377	9	US-10-176-918-361 Sequence 361, App
13	1377	100.0	1377	9	US-10-137-865-361 Sequence 361, App
14	1377	100.0	1377	9	US-10-137-865-361 Sequence 361, App
15	1377	100.0	1377	9	US-10-140-474-361 Sequence 361, App
16	1377	100.0	1377	9	US-10-142-431-361 Sequence 361, App
17	1377	100.0	1377	9	US-10-143-114-361 Sequence 361, App
18	1377	100.0	1377	9	US-10-140-002-361 Sequence 361, App
19	1377	100.0	1377	9	US-10-142-419-361 Sequence 361, App

20	1377	100.0	1377	9	US-10-123-262-361 Sequence 361, App
21	1377	100.0	1377	9	US-10-142-423-361 Sequence 361, App
22	1377	100.0	1377	9	US-10-121-050-361 Sequence 361, App
23	1377	100.0	1377	9	US-10-141-755-361 Sequence 361, App
24	1377	100.0	1377	9	US-10-143-023-361 Sequence 361, App
25	1377	100.0	1377	9	US-10-123-108-361 Sequence 361, App
26	1377	100.0	1377	9	US-10-123-236-361 Sequence 361, App
27	1377	100.0	1377	9	US-10-123-261-361 Sequence 361, App
28	1377	100.0	1377	9	US-10-140-921-361 Sequence 361, App
29	1377	100.0	1377	9	US-10-140-928-361 Sequence 361, App
30	1377	100.0	1377	9	US-10-121-045-361 Sequence 361, App
31	1377	100.0	1377	9	US-10-123-292-361 Sequence 361, App
32	1377	100.0	1377	9	US-10-123-903-361 Sequence 361, App
33	1377	100.0	1377	9	US-10-124-819-361 Sequence 361, App
34	1377	100.0	1377	9	US-10-124-823-361 Sequence 361, App
35	1377	100.0	1377	9	US-10-140-925-361 Sequence 361, App
36	1377	100.0	1377	9	US-10-160-498-361 Sequence 361, App
37	1377	100.0	1377	9	US-09-944-884-41 Sequence 41, App1
38	1377	100.0	1377	9	US-10-121-041-361 Sequence 361, App
39	1377	100.0	1377	9	US-10-121-043-361 Sequence 361, App
40	1377	100.0	1377	9	US-10-121-047-361 Sequence 361, App
41	1377	100.0	1377	9	US-10-123-215-361 Sequence 361, App
42	1377	100.0	1377	9	US-10-123-902-361 Sequence 361, App
43	1377	100.0	1377	9	US-10-123-908-361 Sequence 361, App
44	1377	100.0	1377	9	US-10-123-909-361 Sequence 361, App
45	1377	100.0	1377	9	US-10-123-910-361 Sequence 361, App

ALIGNMENTS

RESULT 1
US-09-944-413-41
Sequence 41, Application US/09944413
Patent No. US20020155600A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944, 413
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866, 028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067, 411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069, 334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069, 696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 694
PRIOR FILING DATE: December 16, 1997


```

: PRIOR APPLICATION NUMBER: 60/069,702
: PRIOR FILING DATE: December 16, 1997
: PRIOR APPLICATION NUMBER: 60/069,870
: PRIOR FILING DATE: December 17, 1997
: PRIOR APPLICATION NUMBER: 60/069,873
: PRIOR FILING DATE: December 17, 1997
: PRIOR APPLICATION NUMBER: 60/068,017
: PRIOR FILING DATE: December 18, 1997
: PRIOR APPLICATION NUMBER: 60/070,440
: PRIOR FILING DATE: January 5, 1998
: PRIOR APPLICATION NUMBER: 60/074,086
: PRIOR FILING DATE: February 9, 1998
: PRIOR APPLICATION NUMBER: 60/074,092
: PRIOR FILING DATE: February 9, 1998
: PRIOR APPLICATION NUMBER: 60/075,945
: PRIOR FILING DATE: February 25, 1998
: PRIOR APPLICATION NUMBER: 60/112,850
: PRIOR FILING DATE: December 16, 1998
: PRIOR APPLICATION NUMBER: 60/113,296
: PRIOR FILING DATE: December 22, 1998
: PRIOR APPLICATION NUMBER: 60/146,222
: PRIOR FILING DATE: July 28, 1999
: PRIOR APPLICATION NUMBER: PCT/US98/19330
: PRIOR FILING DATE: September 16, 1998
: PRIOR APPLICATION NUMBER: PCT/US98/25108
: PRIOR FILING DATE: December 1, 1998
: PRIOR APPLICATION NUMBER: 09/216,021
: PRIOR FILING DATE: December 16, 1998
: PRIOR APPLICATION NUMBER: 09/218,517
: PRIOR FILING DATE: December 22, 1998
: PRIOR APPLICATION NUMBER: 09/254,311
: PRIOR FILING DATE: March 3, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: June 22, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIOR FILING DATE: September 15, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/28409
: PRIOR FILING DATE: No. US200201560041amember 30, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/28313
: PRIOR FILING DATE: No. US200201560041amember 30, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/28301
: PRIOR FILING DATE: December 1, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/30095
: PRIOR FILING DATE: December 16, 1999
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: February 11, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: February 22, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: March 2, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: March 30, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/14042
: PRIOR FILING DATE: May 22, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/20710
: PRIOR FILING DATE: July 28, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: December 1, 2000
: PRIOR APPLICATION NUMBER: PCT/US01/06520
: PRIOR FILING DATE: February 28, 2001
: NUMBER OF SEQ ID NOS: 120
: SEQ ID NO: 41
: LENGTH: 1377
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-09-944-413-41

```

```

Query Match      100.0%; Score 1377; DB 9; length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACTGTTCTCTTGGAGTCTGGAGAGGAACGGGACCGCCGACGAGCCAACTCAGAC 60
|||||

```

D	1	GACTAGTTCCTTGGAGCTCTGGAGGAGGAAAGCGAGCGCGCAGGGAGCGAACCAAGAC	60
Q	61	TGGGGTGAACCGCAGAGCGAGGGGGGCTCTGGCCGGGGAGAAAGCGGGGGCTGAGACACA	120
D	61	TGGGGTGAACCGCAGAGCGAGGGGGGCGCTTGGCCGGGGAGAAAGCGGGGGCTGAGACACA	120
Q	121	CCAACTGAAGAGGTCTCGAGATGACGAGAGCGCCCGCAAGAGAGCCATCGGGAGAGCCGGAGAGG	180
D	121	CCAACTGAAGAGGTCTCGAGATGAGCCAGCGCCCGCAAGAGAGCCATCGGGAGAGCCGGAGAGG	180
Q	181	GGAATCGAGAGGAGCCCTCGAGCTCCGGAGCTCCGGTGCAGAGCGCTATGAGGCCACTCT	240
D	181	GGAATCGAGAGGAGCCCTCGAGCTCCGGAGCTCCGGTGCAGAGCGCTATGAGGCCACTCT	240
Q	241	CGTCTGTGTCCTCGGGGGCTGGGGGGCGAGCTGCGCCCACTGGACACAAAGATCCC	300
D	241	CGTCTGTGTCCTCGGGGGCTGGGGGGCGAGCTGCGCCCACTGGACACAAAGATCCC	300
Q	301	CAGCCTCTGCCCCGGGGAGACCCCGGCTTTCAGAGCAGCCGGGACCAATGACAGAGG	360
D	301	CAGCCTCTGCCCCGGGGAGACCCCGGCTTTCAGAGCAGCCGGGACCAATGACAGAGG	360
Q	361	CTTTCGCGGGCGCGATGGCGCGCAGCGGCGGAGCGCGCTCCGGGGCTCCGGAGAGAA	420
D	361	CTTTCGCGGGCGCGCATGGCGCGCAGCGGCGGAGCGGCGCTCCGGGGCTCCGGAGAGAA	420
Q	421	AGGCGAGGGCGGAGGCGCGGGACCTGCGGGAGCCTGAGGGGAGACCCCGGGGCGGAGGAGA	480
D	421	AGGCGAGGGCGGAGGCGCGGGAGCTGCGGGAGCCTGAGGGGAGACCCCGGGGCGGAGGAGA	480
Q	481	GAGCGGAAACCGCGGGGGCCACCGGGCTGCGGGAGATGCTCGGTGCCTCGCATCGC	540
D	481	GAGCGGAAACCGCGGGGGCCACCGGGCTGCGGGAGATGCTCGGTGCCTCGCATCGC	540
Q	541	CTTCAGGGCCAAAGCGCTCCAGAGAGCGGGTGCCTCGGCGGCTGACACCTTGGCCCTT	600
D	541	CTTCAGGGCCAAAGCGCTCCAGAGAGCGGGTGCCTCGGCGGCTGACACCTTGGCCCTT	600
Q	601	CGACCGGCTGTGTGACAGAGCAGGGACATTAAGAGCGGTCAACCGGCAAGTTCACTG	660
D	601	CGACCGGCTGTGTGACAGAGCAGGGACATTAAGAGCGGTCAACCGGCAAGTTCACTG	660
Q	661	CCAGGTGCTGGGGTCTACTACTTCCGCCGTTCATGCAACGGTCTACCGGGCAGCCTGCA	720
D	661	CCAGGTGCTGGGGTCTACTACTTCCGCCGTTCATGCAACGGTCTACCGGGCAGCCTGCA	720
Q	721	GTTTGATCTGGTGAAGATGGCGAATCATATGCGCTCTTCTTCAGATTCTTCCGGGGGTG	780
D	721	GTTTGATCTGGTGAAGATGGCGAATCATATGCGCTCTTCTTTCAGATTCTTCCGGGGGTG	780
Q	781	GCCCAAGCGAGCTCGCTCTCGGGGGGGGCGATGATGAGCGTGGAGCCTTAGAGCAAGT	840
D	781	GCCCAAGCGAGCTCGCTCTCGGGGGGGGCGATGATGAGCGTGGAGCCTTAGAGCAAGT	840
Q	841	GTTGGTGAAGGTGGGTGTGGGTGACTACTATTTGGCATCTATGCAAGCATCAAGACAGAG	900
D	841	GTTGGTGAAGGTGGGTGTGGGTGACTACTATTTGGCATCTATGCAAGCATCAAGACAGAG	900
Q	901	CACCTTCTCCGGATTTCTGTGTACTCCGAGATGGCAAGCTCCCAAGTCTTTGCTTAGTG	960
D	901	CACCTTCTCCGGATTTCTGTGTACTCCGAGATGGCAAGCTCCCAAGTCTTTGCTTAGTG	960
Q	961	CCCACTGCAAAATGATGCTCATATGCTCTCACTCTAGAAAGAGGGGTGTAAGGTGACAAACA	1020
D	961	CCCACTGCAAAATGATGCTCATATGCTCTCACTCTAGAAAGAGGGGTGTAAGGTGACAAACA	1020
Q	1021	GGTGATCCAGAGGGCTGGGCCCCCTCGAATAATTGTAAATGACTAGGAGGTGGGTAGA	1080
D	1021	GGTGATCCAGAGGGGTGGGCCCCCTCGAATAATTGTAAATGACTAGGAGGTGGGTAGA	1080
Q	1081	GCACTTTCCTGCTCTGCTGTGGCAAGAAATGGAAACATGAGCTGTCTGCGATCAGGCTG	1140
D	1081	GCACTTTCCTGCTCTGCTGTGGCAAGAAATGGAAACATGAGCTGTCTGCGATCAGGCTG	1140

[illegible]

```

1      RESULT 2
2      US-09-944-403-41
3      ; Sequence 41, Application US/09944403
4      ; Patent No. US2002016513A1
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Baker, Kevin
7      ; APPLICANT: Botstein, David
8      ; APPLICANT: Eaton, Dan
9      ; APPLICANT: Ferrara, Napoleone
10     ; APPLICANT: Filvaroff, Ellen
11     ; APPLICANT: Gerritsen, Mary
12     ; APPLICANT: Goddard, Audrey
13     ; APPLICANT: Godowski, Paul
14     ; APPLICANT: Grimaldi, Christopher
15     ; APPLICANT: Gurney, Austin
16     ; APPLICANT: Hillan, Kenneth
17     ; APPLICANT: Kljavin, Ivar
18     ; APPLICANT: Napier, Mary
19     ; APPLICANT: Roy, Margaret
20     ; APPLICANT: Tumas, Daniel
21     ; APPLICANT: Wood, William
22     ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
23     ; FILE OF INVENTION: ACIDS ENCODING THE SAME
24     ; FILE REFERENCE: P2548P1C1
25     ; CURRENT APPLICATION NUMBER: US/09/944,403
26     ; CURRENT FILING DATE: 2001-09-26
27     ; PRIOR APPLICATION NUMBER: 09/866,028
28     ; PRIOR FILING DATE: 2001-05-25
29     ; PRIOR APPLICATION NUMBER: 60/067,411
30     ; PRIOR FILING DATE: December 3, 1997
31     ; PRIOR APPLICATION NUMBER: 60/069,334
32     ; PRIOR FILING DATE: December 11, 1997
33     ; PRIOR APPLICATION NUMBER: 60/069335
34     ; PRIOR FILING DATE: December 11, 1997
35     ; PRIOR APPLICATION NUMBER: 60/069,278
36     ; PRIOR FILING DATE: December 11, 1997
37     ; PRIOR APPLICATION NUMBER: 60/069,425
38     ; PRIOR FILING DATE: December 12, 1997
39     ; PRIOR APPLICATION NUMBER: 60/069,686
40     ; PRIOR FILING DATE: December 16, 1997
41     ; PRIOR APPLICATION NUMBER: 60/069,694
42     ; PRIOR FILING DATE: December 16, 1997
43     ; PRIOR APPLICATION NUMBER: 60/069,702
44     ; PRIOR FILING DATE: December 16, 1997
45     ; PRIOR APPLICATION NUMBER: 60/069,870
46     ; PRIOR FILING DATE: December 17, 1997
47     ; PRIOR APPLICATION NUMBER: 60/069,873
48     ; PRIOR FILING DATE: December 17, 1997
49     ; PRIOR APPLICATION NUMBER: 60/069,017
50     ; PRIOR FILING DATE: December 18, 1997
51     ; PRIOR APPLICATION NUMBER: 60/070,440
52     ; PRIOR FILING DATE: January 5, 1998
53     ; PRIOR APPLICATION NUMBER: 60/074,086
54     ; PRIOR FILING DATE: February 9, 1998
55     ; PRIOR APPLICATION NUMBER: 60/074,092

```

```

1 PRIOR FILING DATE: February 9, 1998
2 PRIOR APPLICATION NUMBER: 60/075,945
3 PRIOR FILING DATE: February 25, 1998
4 PRIOR APPLICATION NUMBER: 60/112,850
5 PRIOR FILING DATE: December 16, 1998
6 PRIOR APPLICATION NUMBER: 60/111,296
7 PRIOR FILING DATE: December 22, 1998
8 PRIOR APPLICATION NUMBER: 60/146,222
9 PRIOR FILING DATE: July 28, 1999
10 PRIOR APPLICATION NUMBER: PCT/US98/19330
11 PRIOR FILING DATE: September 16, 1998
12 PRIOR APPLICATION NUMBER: PCT/US98/25108
13 PRIOR FILING DATE: December 1, 1998
14 PRIOR APPLICATION NUMBER: 09/216,021
15 PRIOR FILING DATE: December 16, 1998
16 PRIOR APPLICATION NUMBER: 09/218,517
17 PRIOR FILING DATE: December 22, 1998
18 PRIOR APPLICATION NUMBER: 09/254,311
19 PRIOR FILING DATE: March 3, 1999
20 PRIOR APPLICATION NUMBER: PCT/US99/12252
21 PRIOR FILING DATE: June 22, 1999
22 PRIOR APPLICATION NUMBER: PCT/US99/21090
23 PRIOR FILING DATE: September 15, 1999
24 PRIOR APPLICATION NUMBER: PCT/US99/28409
25 PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
26 PRIOR APPLICATION NUMBER: PCT/US99/28313
27 PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
28 PRIOR APPLICATION NUMBER: PCT/US99/28301
29 PRIOR FILING DATE: December1, 1999
30 PRIOR APPLICATION NUMBER: PCT/US99/30095
31 PRIOR FILING DATE: December 16, 1999
32 PRIOR APPLICATION NUMBER: PCT/US00/03555
33 PRIOR FILING DATE: February 11, 2000
34 PRIOR APPLICATION NUMBER: PCT/US00/04414
35 PRIOR FILING DATE: February 22, 2000
36 PRIOR APPLICATION NUMBER: PCT/US00/05841
37 PRIOR FILING DATE: March 2, 2000
38 PRIOR APPLICATION NUMBER: PCT/US00/08449
39 PRIOR FILING DATE: March 30, 2000
40 PRIOR APPLICATION NUMBER: PCT/US00/14042
41 PRIOR FILING DATE: May 22, 2000
42 PRIOR APPLICATION NUMBER: PCT/US00/20710
43 PRIOR FILING DATE: July 28, 2000
44 PRIOR APPLICATION NUMBER: PCT/US00/32678
45 PRIOR FILING DATE: December 1, 2000
46 PRIOR APPLICATION NUMBER: PCT/US01/06520
47 PRIOR FILING DATE: February 28, 2001
48 NUMBER OF SEQ ID NOS: 120
49 SEQ ID NO 41
50 LENGTH: 1377
51 TYPE: DNA
52 ORGANISM: Homo Sapien
53 US-09-944-403-41

```

Query Match	100.0%	Score 1377	DB 9	Length 1377	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 1377	Conservative	0	Mismatches	0	Indels
					Gaps
Qy	1	GACTAGTCTCTTGGAGCTCTGGAGAGAGAAAGCGAGCCGSCAGAGGACGNAACCAAGNC	60		
Db	1	GACTAGTCTCTTGGAGCTCTGGAGAGAGAAAGCGAGCCGSCAGAGGACGNAACCAAGNC	60		
Qy	61	TGGGGTGAACGGACAGGGGCAAGGGGCGCCCTGGCCGGGGAGAAAGCCGGGGGCTTGAGACCA	120		
Db	61	TGGGGTGAACGGACAGGGGCAAGGGGCGCCCTGGCCGGGGAGAAAGCCGGGGGCTTGAGACCA	120		
Qy	121	CCAACTGAGAGGCTCGAGAGTAGACGACCCCCGAGAAAGAGCCCATCGAGGAGCCCGGAGGG	180		
Db	121	CCAACTGAGAGGCTCGAGAGTAGACGACCCCCGAGAAAGAGCCCATCGAGGAGCCCGGAGGG	180		
Qy	181	GAGGACTGCGAGAGNACCCCGCGTCTCCGGGTCTCCAGCGCTTAGAGGCACTCTT	240		
Db	181	GAGGACTGCGAGAGNACCCCGCGTCTCCGGGTCTCCAGCGCTTAGAGGCACTCTT	240		

QY 661 CCAAGTGTGGGGTCTACTACTTGGCGTCCAGTCCAGCGGTCCAGCGCTGCA 720
Db 661 CCAAGTGTGGGGTCTACTACTTGGCGTCCAGTCCAGCGGTCCAGCGCTGCA 720
QY 721 GTTGTATCTGTGAAGATGCGCATTCATTCCTCTTTCAGTTTTCGCGGGGTG 780
Db 721 GTTGTATCTGTGAAGATGCGCATTCATTCCTCTTTCAGTTTTCGCGGGGTG 780
QY 781 GCGCCAGCCAGCTCTCTCTCGGGGGGGCCATGTGTAGAGCTGAGCCCTGAGCAAGT 840
Db 781 GCGCCAGCCAGCTCTCTCTCGGGGGGGCCATGTGTAGAGCTGAGCCCTGAGCAAGT 840
QY 841 GTGGGTGCAAGTGTGGGTGTGACATTCATTCGATTCATTCGATTCATTCGATTC 900
Db 841 GTGGGTGCAAGTGTGGGTGTGACATTCATTCGATTCATTCGATTCATTCGATTC 900
QY 901 CACCTTCTCCGATTTCTGTGTGTACTCCAGTGTGCAAGCTCCCAAGTCTTTCCTT 960
Db 901 CACCTTCTCCGATTTCTGTGTGTACTCCAGTGTGCAAGCTCCCAAGTCTTTCCTT 960
QY 961 CCACTGTGCAAGTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 961 CCACTGTGCAAGTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1021 GGTTCATCCAGAGGGGTGTGCGCCCGCTGGATTTGTGATGATGATGATGATG 1080
Db 1021 GGTTCATCCAGAGGGGTGTGCGCCCGCTGGATTTGTGATGATGATGATGATG 1080
QY 1081 GCACTCTCCGCT 1140
Db 1081 GCACTCTCCGCT 1140
QY 1141 GCGAGTGTGGGGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1141 GCGAGTGTGGGGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 1201 GTGTAGTGTGGGGT 1260
Db 1201 GTGTAGTGTGGGGT 1260
QY 1261 CTCTGTCTCTCTGT 1320
Db 1261 CTCTGTCTCTCTGT 1320
QY 1321 GATCAGT 1377
Db 1321 GATCAGT 1377

RESULT 5
Sequence 41, Application US/09944907
Publication No. US20020198147A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Baton, Dan
APPLICANT: Perrata, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Geritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944, 907
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866, 028
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 41
LENGTH: 1377
TYPE: DNA
ORGANISM: Homo Sapien
US-09-944-907-41

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACTAGTCTCTTGTAGTCTGTGAGAGAGAAAGCGAGCGGACGAGGACGAAACAGAG 60
Db 1 GACTAGTCTCTTGTAGTCTGTGAGAGAGAAAGCGAGCGGACGAGGACGAAACAGAG 60
QY 61 TGGGTGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 61 TGGGTGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 CCAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Db 121 CCAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 181 GGGAGTGTGAG 240
Db 181 GGGAGTGTGAG 240
QY 241 GGTCTGTGTCTCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 300
Db 241 GGTCTGTGTCTCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 300
QY 301 CAGCCTCTGCGCGGAG 360
Db 301 CAGCCTCTGCGCGGAG 360
QY 361 CTTCGCGGAG 420
Db 361 CTTCGCGGAG 420
QY 421 AGGCGAG 480
Db 421 AGGCGAG 480
QY 481 GCGGAG 540
Db 481 GCGGAG 540
QY 541 CTTCAGCGGAG 600
Db 541 CTTCAGCGGAG 600
QY 601 CCAAGCGGAG 660
Db 601 CCAAGCGGAG 660
QY 661 CCAAGTGTGTGGGTGTGACTTCTGCGCTGTGATGCAAGCGGTCTGACCGGAGCTG 720
Db 661 CCAAGTGTGTGGGTGTGACTTCTGCGCTGTGATGCAAGCGGTCTGACCGGAGCTG 720
QY 721 GTTGTATCTGTGAAGATGCGCATTCATTCCTCTTTCAGTTTTCGCGGGGTG 780
Db 721 GTTGTATCTGTGAAGATGCGCATTCATTCCTCTTTCAGTTTTCGCGGGGTG 780
QY 781 GCGCAAGCGAGCTGTGTGGGGGGGGCCATGTGTAGAGCTGAGCCCTGAGCAAGT 840
Db 781 GCGCAAGCGAGCTGTGTGGGGGGGGCCATGTGTAGAGCTGAGCCCTGAGCAAGT 840
QY 841 GTGGGTGCAAGTGTGGGTGTGACATTCATTCGATTCATTCGATTCATTCGATTC 900

Db	841	GTGGGTGAGAGTGGGTGTGGGTGATCTAATTTGGATCTATATCCGAGCATCAAGACAGACAG	900
Qy	901	CACCTTCTCCGGATTTCTGTGTACTCCGACTGCGACAGCTTCCCAAGCTTTTGCTTAAGTG	960
Db	901	CACCTTCTCCGGATTTCTGTGTACTCCGACTGCGACAGCTTCCCAAGCTTTTGCTTAAGTG	960
Qy	961	CCCACTGGAAGTGAAGTCATGAGCTCTCACTCCCTGGAAGAGAGAGGTGAGAGGCTGAACAACA	1020
Db	961	CCCACTGGAAGTGAAGTCATGAGCTCTCACTCCCTGGAAGAGAGAGGTGAGAGGCTGAACAACA	1020
Qy	1021	GGTCAATCAGGAGGGCTGAGCCCCCTCGAATAATTTGTGAATGACTGAGAGGTGGGTGAGA	1080
Db	1021	GGTCAATCAGGAGGGCTGAGCCCCCTCGAATAATTTGTGAATGACTGAGAGGTGGGTGAGA	1080
Qy	1081	GCACTCTCCGTCTGTGCTGTGCGAAGAAATGGGAAACAGTGGCTGTCTGCGAATCAGGTCTG	1140
Db	1081	GCACTCTCCGTCTGTGCTGTGCGAAGAAATGGGAAACAGTGGCTGTCTGCGAATCAGGTCTG	1140
Qy	1141	GCAGATATGGGACAGTGGCTGGAATTTCTGCCCAAGCCAGAGAGATGTCTGTCTGTGCGAA	1200
Db	1141	GCAGATATGGGACAGTGGCTGGAATTTCTGCCCAAGCCAGAGAGATGTCTGTCTGTGCGAA	1200
Qy	1201	GTGTAAAGTCCCCCAAGTTGCTGTGTGTCAGAGACCCAGTGGAGGTCTCTCTTCCTGGTTC	1260
Db	1201	GTGTAAAGTCCCCCAAGTTGCTGTGTGTCAGAGACCCAGTGGAGGTCTCTCTTCCTGGTTC	1260
Qy	1261	CTGTGCTTCTCTGAATCTCTCCCAAGCCCTCTGTGCTCTCGGAGGCGGAGCCCTTTTCTCAGA	1320
Db	1261	CTGTGCTTCTCTGAATCTCTCCCAAGCCCTCTGTGCTCTCGGAGGCGGAGCCCTTTTCTCAGA	1320
Qy	1321	GATCACTCAATTAACCTTAAGAACCTCAATTAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1377
Db	1321	GATCACTCAATTAACCTTAAGAACCTCAATTAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1377

RESULT 6

US-09-944-929-41
; Sequence 41, Application US/09944929
; Publication No. US20020197612A1

```

/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin
/ APPLICANT: Batstein, David
/ APPLICANT: Batson, Dan
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gerritsen, Mary
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul
/ APPLICANT: Grimaldi, Christopher
/ APPLICANT: Gurney, Austin
/ APPLICANT: Hillman, Kenneth
/ APPLICANT: Kijavini, Ivar
/ APPLICANT: Napier, Mary
/ APPLICANT: Roy, Margaret
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Wood, William
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P2548P1C1
/ CURRENT APPLICATION NUMBER: US/09/944,929
/ CURRENT FILING DATE: 2001-08-31
/ PRIOR APPLICATION NUMBER: 09/866,028
/ PRIOR FILING DATE: 2001-05-25
/ NUMBER OF SEQ ID NOS: 120

```

```

; SEQ ID NO 41
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-944-929-41

```

Query Match	Score	DB	Length
100.0%;	1377;	9;	1377;
Best Local Similarity	Pred.	No.	0;

	Matches	1377	Conservative	0	Mismatches	0	Indels	0	Gaps	0
QY	1	GACTAGTTCCTCTTGGAGTCTGGGAGAGGAAAGCCGAGCCGACAGGAGCGCAACCAAGAC	60							
Db	1	GACTAATTCCTTGGAGTCTGGAGGAGGAAACCGAGCCGACAGGAGCGAACCAAGAC	60							
QY	61	TGGGGTGAACGACAGGAGCGAGGAGCGCTTGGCCCGGAGGAGAAAGCGGAGGCTGGAGCAACA	120							
Db	61	TGGGGTGAACGACAGGAGCGAGGAGCGCTTGGCCCGGAGGAGAAAGCGGAGGCTTGGAGCAACA	120							
QY	121	CCAACTGGAGGATCCGAGTGAAGCGCCCGAAGAGGCCATCGGAGAGCCGAGAGCG	180							
Db	121	CCAACTGGAGGATCCGAGTGAAGCGCCCGAAGAGGCCATCGGAGAGCCGAGAGCG	180							
QY	181	GGGACGTGCAAGAGACCCCGGCGCTCCCGGAGCTCCCGTGCACAGCGCTATGAGGCCACTCT	240							
Db	181	GGGACGTGCAAGAGACCCCGGCGCTCCCGGAGCTCCCGTGCACAGCGCTATGAGGCCACTCT	240							
QY	241	CGTCTGCTGCTCTTGGGCTTGGCGGCGCGGCTCGCCCACTGGACGACAAGATCC	300							
Db	241	CGTCTGCTGCTCTTGGGCTTGGCGGCGCGGCTCGCCCACTGGACGACAAGATCC	300							
QY	301	CAGCCTCTGCTCCGAGGACCCCGGCTTCCAGGCAAGCGGAGCCACAATGGCAGCCAGGG	360							
Db	301	CAGCCTCTGCTCCGAGGACCCCGGCTTCCAGGCAAGCGGAGCCACAATGGCAGCCAGGG	360							
QY	361	CTTGGCGGAGCGAGATGGACGAGCGGAGCGGAGCGGAGCGCGCGGAGGCTCCGGAGAGAA	420							
Db	361	CTTGGCGGAGCGGAGATGGACGAGCGGAGCGGAGCGGAGCGCGCGGAGGCTCCGGAGAGAA	420							
QY	421	AGGCGAGAGCGGAGAGCGCGGAGCTGCGGAGACTGAGGAGAACCCCGAGCGGAGAGAA	480							
Db	421	AGGCGAGAGCGGAGAGCGCGGAGCTGCGGAGACTGAGGAGAACCCCGAGCGGAGAGAA	480							
QY	481	GGGAGGAGACCGGAGGAGCCACCGGAGCTGCGGAGAGATGCTCGGAGCTCCGAGATCCG	540							
Db	481	GGGAGGAGACCGGAGGAGCCACCGGAGCTGCGGAGAGATGCTCGGAGCTCCGAGATCCG	540							
QY	541	CTTCAAGGCGCAAGCGCTCCGAGAGCGGAGTGCCTCGCGCTGACGCAACCTTGGCCTT	600							
Db	541	CTTCAAGGCGCAAGCGCTCCGAGAGCGGAGTGCCTCGCGCTGACGCAACCTTGGCCTT	600							
QY	601	CGAACCGGCTGTGTGTGAAGAGAGAGGACATTAAGAGCCGTCAACCGGCAAGTTCACTG	660							
Db	601	CGAACCGGCTGTGTGTGAAGAGAGAGGACATTAAGAGCCGTCAACCGGCAAGTTCACTG	660							
QY	661	CGAGGTGCTGGGGGTCTACTTCTTCCGCTGCAAGCCGTCTAACCGGAGCAACCTGCA	720							
Db	661	CGAGGTGCTGGGGGTCTACTTCTTCCGCTGCAAGCCGTCTAACCGGAGCAACCTGCA	720							
QY	721	GTTTGATCTGTGAAGAGATGAGCGAATCATTGCCTTTCTTCCAGTTTTCGGAGGAGTG	780							
Db	721	GTTTGATCTGTGAAGAGATGAGCGAATCATTGCCTTTCTTCCAGTTTTCGGAGGAGTG	780							
QY	781	GGCCAAAGCCAGCCCTCGCTCCGGAGGAGGACATGTGAGGTGAGGCTGAGCAACAAGT	840							
Db	781	GGCCAAAGCCAGCCCTCGCTCCGGAGGAGGACATGTGAGGTGAGGCTGAGCAACAAGT	840							
QY	841	GTGGGTGACAGGTGGGTGTGGGTGAATCATTTGGCATCTATGCAAGCATCAAGACAG	900							
Db	841	GTGGGTGACAGGTGGGTGTGGGTGAATCATTTGGCATCTATGCAAGCATCAAGACAG	900							
QY	901	CACCTTCTCGGAAATTCGTGGTGTATCTCGACTGGCAAGCTCCCAAGCTTTTGTCTTATG	960							
Db	901	CACCTTCTCGGAAATTCGTGGTGTATCTCGACTGGCAAGCTCCCAAGCTTTTGTCTTATG	960							
QY	961	CCCACTGCAAAAGTGAAGCTCATGCTCTCACTCTTGAAGAGGAGGTGAGGCTGCAACCA	1020							
Db	961	CCCACTGCAAAAGTGAAGCTCATGCTCTCACTCTTGAAGAGGAGGTGAGGCTGCAACCA	1020							
QY	1021	GGTCAATCAGAGGAGCGTCCCGCTGGAATATGTGAATTAATAGGAGGTGGGTGATA	1080							
Db	1021	GGTCAATCAGAGGAGCGTCCCGCTGGAATATGTGAATTAATAGGAGGTGGGTGATA	1080							


```

1 PRIOR FILING DATE: 1997-10-17
2 PRIOR APPLICATION NUMBER: 60/062287
3 PRIOR FILING DATE: 1997-10-17
4 PRIOR APPLICATION NUMBER: 60/062814
5 PRIOR FILING DATE: 1997-10-24
6 PRIOR APPLICATION NUMBER: 60/062816
7 PRIOR FILING DATE: 1997-10-24
8 PRIOR APPLICATION NUMBER: 60/063045
9 PRIOR FILING DATE: 1997-10-24
10 PRIOR APPLICATION NUMBER: 60/063082
11 PRIOR FILING DATE: 1997-10-31
12 PRIOR APPLICATION NUMBER: 60/063127
13 PRIOR FILING DATE: 1997-10-24
14 PRIOR APPLICATION NUMBER: 60/063327
15 PRIOR FILING DATE: 1997-10-27
16 PRIOR APPLICATION NUMBER: 60/063329
17 PRIOR FILING DATE: 1997-10-27
18 PRIOR APPLICATION NUMBER: 60/063550
19 PRIOR FILING DATE: 1997-10-28
20 PRIOR APPLICATION NUMBER: 60/063561
21 PRIOR FILING DATE: 1997-10-28

```

US-10-028-072-361

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Desnoyers, Luc

APPLICANT: Gao, Wei-Qiang

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Sherwood, Steven

APPLICANT: Stewart, Timothy

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

10 TITLE OF INVENTION:

CURRENT APPLICATION NUMBER:

PRIOR APPLICATION NUMBER: 60

PRIOR FILING DATE: 1997-06-11
PRIOR APPLICATION NUMBER: 60

PRIOR FILING DATE: 1997-08-2

PRIOR FILING DATE: 1997-09-11

PRIOR FILING DATE: 1997-09-11

PRIOR APPLICATION NUMBER: 60
PRIOR FILING DATE: 1997-09-1

PRIOR APPLICATION NUMBER: 60

PRIOR APPLICATION NUMBER: 60

PRIOR APPLICATION NUMBER: 60

PRIOR FILING DATE: 1997-09-11
PRIOR APPLICATION NUMBER: 60

PRIOR FILING DATE: 1997-09-11

PRIOR FILING DATE: 1997-09-1

PRIOR FILING DATE: 1997-09-2

PRIOR APPLICATION NUMBER: 60

! PRIOR APPLICATION NUMBER: 60

PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360

PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACTAGTTCTCTGAGTCTGGAGAGAGAAAGCGAGCGGAGGAGGAAACAGAGAC 60
1 GACTAGTTCTCTGAGTCTGGAGAGAGAAAGCGAGCGGAGGAGGAAACAGAGAC 60
1 GACTAGTTCTCTGAGTCTGGAGAGAGAAAGCGAGCGGAGGAGGAAACAGAGAC 60

61 TGAGGTGACGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
61 TGAGGTGACGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120

121 CCAACTGAGAGGTCCGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
121 CCAACTGAGAGGTCCGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180

181 GGAAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
181 GGAAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240

241 GGTCTGCTCTCTCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
241 GGTCTGCTCTCTCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

301 CAGCTCTGCTCTCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
301 CAGCTCTGCTCTCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

361 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
361 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420

421 AGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
421 AGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480

481 GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
481 GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540

541 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
541 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600

541 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
541 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600

601 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
601 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660

601 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
601 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660

661 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
661 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720

721 GTTTGAATCTGTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
721 GTTTGAATCTGTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780

781 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
781 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840

841 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
841 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900

901 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
901 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960

QY 961 CCCACTGCAAAAGTGAAGTCTATGCTCTCACTCTCTAGAGAGAGGTGTGAGGCTTGACCA 1020
DB 961 CCCACTGCAAAAGTGAAGTCTATGCTCTCACTCTCTAGAGAGAGGTGTGAGGCTTGACCA 1020
QY 1021 GGTTCATCCAGAGAGGCTGGGCCCCCTTGAAATATTGGAATGACTAGGAGGTGGGTAGA 1080
DB 1021 GGTTCATCCAGAGAGGCTGGGCCCCCTTGAAATATTGGAATGACTAGGAGGTGGGTAGA 1080
QY 1081 GCACTCTCCGCTCTGCTGCTGCTGCAAGAAATGAGAGAGGTGCTGCTGAGTACAGGCTG 1140
DB 1081 GCACTCTCCGCTCTGCTGCTGCTGCAAGAAATGAGAGAGGTGCTGCTGAGTACAGGCTG 1140
QY 1141 GCAAGCATGGGGAGGTGGCTGGAATTTCTGCCAAGACCAAGAGAGTGTGCTGCTGCA 1200
DB 1141 GCAAGCATGGGGAGGTGGCTGGAATTTCTGCCAAGACCAAGAGAGTGTGCTGCTGCA 1200
QY 1201 GGTGAAGTCCCGAGTTGCTCTGCTGCTCAAGAGCCCAAGGTGGGTGCTCTCTGCTG 1260
DB 1201 GGTGAAGTCCCGAGTTGCTCTGCTGCTCAAGAGCCCAAGGTGGGTGCTCTCTGCTG 1260
QY 1261 CTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 CTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 GATCACTCAATAAAGCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
DB 1321 GATCACTCAATAAAGCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 8

US-10-121-049-361

Sequence 361, Application US/10121049

Publication No. US20030022239A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Inc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gunney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C17
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 361
LENGTH: 1377
TYPE: DNA
ORGANISM: Homo Sapien
US-10-121-049-361

Query Match

100.0%; Score 1377; DB 9; Length 1377;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACTAGTTCCTCTGAGTCTGAG 60
DB 1 GACTAGTTCCTCTGAGTCTGAG 60

QY 61 TGGGCTGACGCGAGCGGCAAGGCGGCTGCGCGGCGGAGAGAGCGCGGCGGCTGAGACCA 120
DB 61 TGGGCTGACGCGAGCGGCAAGGCGGCGGCTGCGCGGCGGAGAGAGCGCGGCGGCTGAGACCA 120
QY 121 CCACTGAGAGGCTCCGAGATGAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 121 CCACTGAGAGGCTCCGAGATGAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 GGGACTGCGAG 240
DB 181 GGGACTGCGAG 240
QY 241 GGTCTGCTGCTCTGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 300
DB 241 GGTCTGCTGCTCTGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 300
QY 301 CAGCTCTGCGGCGGAG 360
DB 301 CAGCTCTGCGGCGGAG 360
QY 361 CTTCGCGGCGGAG 420
DB 361 CTTCGCGGCGGAG 420
QY 421 AGGCGAG 480
DB 421 AGGCGAG 480
QY 481 GGGCGAG 540
DB 481 GGGCGAG 540
QY 541 CTTCGAGCGGAG 600
DB 541 CTTCGAGCGGAG 600
QY 601 CGACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CGACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 CCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 GTTGAATCTGTAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 GTTGAATCTGTAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 GCCCAAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 GCCCAAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 GTGGGTGAGAGTGGGTGGGTGATCAATTTGATGATGATGATGATGATGATGATGATGAT 900
DB 841 GTGGGTGAGAGTGGGTGGGTGATCAATTTGATGATGATGATGATGATGATGATGATGAT 900
QY 901 CACCTCTCCGAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 901 CACCTCTCCGAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 CCCACTGCAAAAGTGAAGTCTATGCTCTCACTCTCTAGAGAGAGGTGTGAGGCTGACCA 1020
DB 961 CCCACTGCAAAAGTGAAGTCTATGCTCTCACTCTCTAGAGAGAGGTGTGAGGCTGACCA 1020
QY 1021 GGTTCATCCAGAGAGGCTGGGCCCCCTTGAAATATTGGAATGACTAGGAGGTGGGTAGA 1080
DB 1021 GGTTCATCCAGAGAGGCTGGGCCCCCTTGAAATATTGGAATGACTAGGAGGTGGGTAGA 1080
QY 1081 GCACTCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1081 GCACTCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 GCAAGCATGGGGAGGTGGCTGGAATTTCTGCCAAGACCAAGAGAGTGTGCTGCTGCA 1200

DB 1141 GCAGCATGCGGCGATGCTGAGATTTCTGCCCAAGACAGAGAGATGCTGCTGCGCA 1200
QY 1201 GTGTAGTCCCGCATGTTGCTGTGTCAGAGAGCCGAGGAGGAGTCTCTTCTGCTGTC 1260
DB 1201 GTGTAGTCCCGCATGTTGCTGTGTCAGAGAGCCGAGGAGGAGTCTCTTCTGCTGTC 1260
QY 1261 CTCTGCTTCTGTGATCTTCCCAAGCCCTCTGCTCTGAGGAGCGGAGCTTTTCTGAGA 1320
DB 1261 CTCTGCTTCTGTGATCTTCCCAAGCCCTCTGCTCTGAGGAGCGGAGCTTTTCTGAGA 1320
QY 1321 GATCACTCAATAAAGCTTAAGACCTCTATTAATAAAAAAAAAAAAAAAAAAAAAA 1377
DB 1321 GATCACTCAATAAAGCTTAAGACCTCTATTAATAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 9

US-10-123-904-361
; Sequence 361, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1CS4
; CURRENT APPLICATION NUMBER: US/10/123,904
; PRIORITY FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-904-361

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACTAGTCTCTTGGAGTCTGGAGAGAGAAAGCCGAGCCGAGAGAGCAACCAAGAG 60
DB 1 GACTAGTCTCTTGGAGTCTGGAGAGAGAAAGCCGAGCCGAGAGAGCAACCAAGAG 60
QY 61 TGGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 61 TGGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 CCAACTGAGAGGAGTCCGAGATGAGAGAGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 121 CCAACTGAGAGGAGTCCGAGATGAGAGAGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 181 GGGAGCTGCGAG 240
DB 181 GGGAGCTGCGAG 240
QY 241 GGTCTGCTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 241 GGTCTGCTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300

DB 241 GGTCTGCTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 301 CAGCTCTGCT 360
DB 301 CAGCTCTGCT 360
QY 361 CTGTGCT 420
DB 361 CTGTGCT 420
QY 421 AGGCGAGAG 480
DB 421 AGGCGAGAG 480
QY 481 GGGCGAGAG 540
DB 481 GGGCGAGAG 540
QY 541 CTTCAGAG 600
DB 541 CTTCAGAG 600
QY 601 CGAGCGGAG 660
DB 601 CGAGCGGAG 660
QY 661 CGAG 720
DB 661 CGAG 720
QY 721 GTTGTATCTGTGAG 780
DB 721 GTTGTATCTGTGAG 780
QY 781 GCCCAAGCAGGAG 840
DB 781 GCCCAAGCAGGAG 840
QY 841 GTGGGTGAG 900
DB 841 GTGGGTGAG 900
QY 901 CACCTTCTCGGAGTCTGCT 960
DB 901 CACCTTCTCGGAGTCTGCT 960
QY 961 CCACTGCAAGAG 1020
DB 961 CCACTGCAAGAG 1020
QY 1021 GGTATCTCAGAGAG 1080
DB 1021 GGTATCTCAGAGAG 1080
QY 1081 GCACTCTCGCT 1140
DB 1081 GCACTCTCGCT 1140
QY 1141 GCAGCATGCGGAG 1200
DB 1141 GCAGCATGCGGAG 1200
QY 1201 GTGTAGTCCCGCATGTTGCTGTGTCAGAGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 GTGTAGTCCCGCATGTTGCTGTGTCAGAGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 CTCTGCTTCTGTGATCTTCCCAAGCCCTCTGCTCTGAGGAGCGGAGCTTTTCTGAGA 1320
DB 1261 CTCTGCTTCTGTGATCTTCCCAAGCCCTCTGCTCTGAGGAGCGGAGCTTTTCTGAGA 1320
QY 1321 GATCACTCAATAAAGCTTAAGACCTCTATTAATAAAAAAAAAAAAAAAAAAAAAA 1377
DB 1321 GATCACTCAATAAAGCTTAAGACCTCTATTAATAAAAAAAAAAAAAAAAAAAAAA 1377


```

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACTS ENCODING THE SAME
; FILE REFERENCE: P330R1C353
; CURRENT APPLICATION NUMBER: US/10/176,746
; PRIOR FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-746-361

```

```

Query Match      100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GACTAGTCTCTTGAGAGTCTGGAGAGAGAAAGCGGAGCGGAGGAGCGAGACAGAGAC 60
DB 1 GACTAGTCTCTTGAGAGTCTGGAGAGAGAAAGCGGAGCGGAGGAGCGAGACAGAGAC 60
QY 61 TGGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 61 TGGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 CCAACTGAGAGGAGTCCGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 121 CCAACTGAGAGGAGTCCGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 181 GGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 GGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 GGTCTGCTGCTCTGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 241 GGTCTGCTGCTCTGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 301 CAGGCTTGTGCTGCTGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 301 CAGGCTTGTGCTGCTGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 CTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 361 CTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 421 AGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 421 AGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 GGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 481 GGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 541 CTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 CTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 601 CGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 CGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 661 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

```

```

DB 661 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 721 GTTGTATCTGTGAAGAGATGAGGAGATCATGCTCTCTTCTTCCAGATTTTTCGGAGGAGT 780
DB 721 GTTGTATCTGTGAAGAGATGAGGAGATCATGCTCTCTTCTTCCAGATTTTTCGGAGGAGT 780
QY 781 GCTCAAGCCAGGCTCGCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 781 GCTCAAGCCAGGCTCGCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 GTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 GTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 CACTTCTCCGAGATTTCTGAGTACTCCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 901 CACTTCTCCGAGATTTCTGAGTACTCCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
QY 961 CCACTGCAAAAGTGAAGTCACTCTCTCACTCTCTCACTCTCTCACTCTCTCACTCTCTCA 1020
DB 961 CCACTGCAAAAGTGAAGTCACTCTCTCACTCTCTCACTCTCTCACTCTCTCACTCTCTCA 1020
QY 1021 GGTATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GGTATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GCACTCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 GCACTCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 GCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 GCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 GTGTAAGTCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 GTGTAAGTCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 CTCTGCTTCTGTGATCTCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1261 CTCTGCTTCTGTGATCTCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1321 GATCATCTCATTAACCTTAAGAACCTCTAATAAATAAATAAATAAATAAATAAATAAATAA 1377
DB 1321 GATCATCTCATTAACCTTAAGAACCTCTAATAAATAAATAAATAAATAAATAAATAAATAA 1377

```

```

RESULT 12
US-10-176-918-361
; Sequence 361, Application US/10176918
; Publication No. US2003002275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918

```


Db	421	AGGGAGAGGACCGAGAGGCCGGGACTGCGCGGACCTTCAGAGGGAGACCCCGAGGCGCAGAGAGA	480
Qy	481	GGCGGAGACCCCGGAGGACCAACCGGAGCTGCGGGGAGAGTGCTCGAGTGCCTCCGAGATCGC	540
Db	481	GGCGGAGACCCCGGAGGACCAACCGGAGCTGCGGGGAGAGTGCTCGAGTGCCTCCGAGATCGC	540
Qy	541	CTTCAGCGCCAAAGCGCTCCAGAGACCGGGGTCCTCGCGCTCTGAGACCACTTCGACCTT	600
Db	541	CTTCAGCGCCAAAGCGCTCCAGAGACCGGGGTCCTCGCGCTCTGAGACCACTTCGACCTT	600
Qy	601	CGACCGGCTGTGTGTGAACGAGCAGGGGACATTAAGAAGCGCTGTACCGGCAATTCACCTG	660
Db	601	CGACCGGCTGTGTGTGAACGAGCAGGGGACATTAAGAAGCGCTGTACCGGCAATTCACCTG	660
Qy	661	CCAGGTGCTGGGGGTCTACTACTTCGCGGTGCATAGCACCCTGTACCGGGGACGCTGCA	720
Db	661	CCAGGTGCTGGGGGTCTACTACTTCGCGGTGCATAGCACCCTGTACCGGGGACGCTGCA	720
Qy	721	GTTTGATCTGTGTGAAGAAGGCGAATCCATTCCTCTTCTTCGAGTTTTCGGGGGGTG	780
Db	721	GTTTGATCTGTGTGAAGAAGGCGAATCCATTCCTCTTCTTCGAGTTTTCGGGGGGTG	780
Qy	781	GCCCCAGCGAGCCCTCGCTTCGGGGGGGGCCCATGCTAGCGCTGAGAGCTGAGACCAAGT	840
Db	781	GCCCCAGCGAGCCCTCGCTTCGGGGGGGGCCCATGCTAGCGCTGAGAGCTGAGACCAAGT	840
Qy	841	GTCGGTGCAGGTGCGGTGTGGGTGACTACATTTGGCATCTATGCGACGATCAAGACAGCAG	900
Db	841	GTCGGTGCAGGTGCGGTGTGGGTGACTACATTTGGCATCTATGCGACGATCAAGACAGCAG	900
Qy	901	CACCTTCTCCGGATTTTCTGTGTACTCCGACTGGGCAAGCTCCCGAGTCTTTGCTTAGTG	960
Db	901	CACCTTCTCCGGATTTTCTGTGTACTCCGACTGGGCAAGCTCCCGAGTCTTTGCTTAGTG	960
Qy	961	CCCACTGCAAAAGTAGCTCATGCTCTCACTCTTGAGAGAGGGGTAGAGGTGACAAACA	1020
Db	961	CCCACTGCAAAAGTAGCTCATGCTCTCACTCTTGAGAGAGGGGTAGAGGTGACAAACA	1020
Qy	1021	GCTCATCAGAGAGGAGCTGGCCCCCTGGAAATATTTGTAATGACTAGGAGAGGTGGGTAGA	1080
Db	1021	GCTCATCAGAGAGGAGCTGGCCCCCTGGAAATATTTGTAATGACTAGGAGAGGTGGGTAGA	1080
Qy	1081	GCACTCTCCGCTCCGCTGCTGGGCAAGAAAGAGTGTCTGTCTGCGATCAGAGTCTG	1140
Db	1081	GCACTCTCCGCTCCGCTGCTGGGCAAGAAAGAGTGTCTGTCTGCGATCAGAGTCTG	1140
Qy	1141	GCAAGCATGAGGACAGTGGCTGATTTCTGCGCAACACAGAGAGTGTCTGTCTGGCAA	1200
Db	1141	GCAAGCATGAGGACAGTGGCTGATTTCTGCGCAACACAGAGAGTGTCTGTCTGGCAA	1200
Qy	1201	GTCGTAATCCCGCAGTTGCTCTGTGTCCAGAGGCCACGCTGGGGGTCTCTCTTCTGTGTC	1260
Db	1201	GTCGTAATCCCGCAGTTGCTCTGTGTCCAGAGGCCACGCTGGGGGTCTCTCTTCTGTGTC	1260
Qy	1261	CTCTGCTTCTCTGTGATCTCTCCCAACCCCTCTCTGTCTGTGGAGCGGCGCTTTTCTGAGA	1320
Db	1261	CTCTGCTTCTCTGTGATCTCTCCCAACCCCTCTCTGTCTGTGGAGCGGCGCTTTTCTGAGA	1320
Qy	1321	GATCATCTCAATTAACCTTAAGAACCTCTCATTAATAAAAAAAAAAAAAAAAAAAAAA	1377
Db	1321	GATCATCTCAATTAACCTTAAGAACCTCTCATTAATAAAAAAAAAAAAAAAAAAAAAA	1377

Search completed: June 20, 2003, 11:21:48
Job time : 228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 06:26:42 ; Search time 1901 Seconds

(without alignments)
11731.292 Million cell updates/sec

Title: US-09-944-403-41

Perfect score: 1377
Sequence: 1 gactagctctctctgagctc.....aaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: 'gb_ges:*
18: em_ges_hum:*
19: em_ges_inv:*
20: em_ges_pln:*
21: em_ges_vrt:*
22: em_ges_fun:*
23: em_ges_mam:*
24: em_ges_mus:*
25: em_ges_other:*
26: em_ges_pro:*
27: em_ges_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	767.8	55.8	869 13	BI763193 603049929
2	750.6	54.5	1154 13	BM547549 AGENCOURT
3	732.8	53.2	1093 14	BM924569 AGENCOURT
4	730.2	53.0	846 13	BI458455 603198837
5	712.6	51.8	801 13	BI490880 603031867
6	672.2	48.8	724 13	BI771879 603055280

7	671	48.7	1045 14	BM920874
8	660.4	48.0	792 13	BI770921
9	641.8	46.6	887 13	BI490062
10	631	45.8	1623 14	BM926477
11	607.2	44.1	630 12	BM970609
12	596	43.3	596 14	BM893587
13	586.4	42.6	596 14	BM544255
14	583.2	42.4	941 13	BI821899
15	578	42.0	578 14	BM893691
16	577.2	41.9	879 13	BI820945
17	572.4	41.6	574 13	BM021127
18	554	40.2	870 9	AF451167
19	551.4	40.0	553 12	BF882978
20	532	38.6	532 12	BF724241
21	521	37.8	521 14	BM966810
22	517	37.5	517 14	BM966954
23	507.8	36.9	574 10	AM150070
24	507.6	36.9	844 10	BE309370
25	489.4	35.5	567 13	BM021381
26	479	34.8	778 12	BF098614
27	463.4	33.7	867 13	BI818537
28	457.4	33.2	1349 9	BM673882
29	453.2	32.9	702 12	BF053285
30	452.4	32.9	454 9	AI805087
31	448.4	32.6	462 10	AM070344
32	440.2	32.0	445 14	BM673416
33	438.4	31.8	842 13	BI688017
34	426.4	31.0	449 13	BM668611
35	424.4	30.8	442 14	BM693336
36	423.4	30.7	548 12	BF906443
37	421.6	30.6	630 12	BF046067
38	409.8	29.8	413 9	AI083823
39	408.4	29.7	410 14	BM707291
40	405.4	29.4	408 9	AI085548
41	402.4	29.2	405 9	AI083824
42	395.6	28.7	547 12	BE722922
43	394.6	28.7	936 12	BF540263
44	392	28.5	392 10	BE138819
45	387.2	28.1	392 10	AM192716

ALIGNMENTS

RESULT 1
BI763193 869 bp mRNA linear EST 25-SEP-2001
LOCUS 603049929F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190222 5'
DEFINITION mRNA sequence.
ACCESSION BI763193
VERSION BI763193.1 GI:15754771
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hautoleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC <http://mgi.mci.nih.gov/>.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: L14M1475 row: 1 column: 07
High quality sequence stop: 777.
Location/Qualifiers
1. 869

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NMGC:5190222"
/clone_1fb="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gubler (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

```

Query Match	55.8%	Score 767.8	DB 13	Length 869
Best Local Similarity	97.7%	Pred. No. 7.9e-96		
Matches	853	Conservative	0	Mismatches 12; Indels 8; Gaps 7
Qy	11	CTTGGAGCTGGGAGAGAGAAAGCCGAGCCGAGCGAGCGAACCAGGACTGGGGGTGACG	70	
Db	1	CTTGGAGCTGGGAGAGAGAAAGCCGAGCCGAGCGAGCGAACCAGGACTGGGGGTGAGG	60	
Qy	71	GCAGGGCAGGGGGCCGCTTGGCCGGGGAGAAAGCCGGGGGCTGGAGCACCACCACTTGAG	130	
Db	61	GCAGGGCAGGTGGCCCTTGGCCGGGGAGAAAGCCGGGGGCTTGAGCACCACCACTTGAG	120	
Qy	131	GATCCGAGATGATACCGAGCCGCCCGAAGAGGCCATATGGGGAGGCCGGGAGGGGGACCTGGCA	190	
Db	121	GATCCGAGATGATACCGAGCCGCCCGAAGAGGCCCATATGGGGAGGCCGGGAGGGGGACCTGGCA	180	
Qy	191	GAGGACCCCGGGGTCCGGGCTCCCGGTGCGAGCGATATGAGGCCACTCTCGTCTGTCTG	250	
Db	181	GAGGACCCCGGGGTCCGGGCTCCCGGTGCGAGCGATATGAGGCCACTCTCGTCTGTCTG	240	
Qy	251	CTCTTGGGCTTGGCGGCGCGGCTTCCGCCCATTTGAGCAGACAAAGATCCCAAGCTTTCG	310	
Db	241	CTCTTGGGCTTGGCGGCGCGGCTTCCGCCCATTTGAGCAGACAAAGATCCCAAGCTTTCG	300	
Qy	311	CCGGGGCACCCTCGGCTTTCAGAGCAGCGCGGGCCACATATGGGACGCCAGGCTTTCGGGCG	370	
Db	301	CCGGGGCACCCTCGGCTTTCAGAGCAGCGCGGGCCACATATGGGAGCCA-GGCTTTCGGGCG	359	
Qy	371	CGCGATGGCCGGGACGCGCCCGGACCGCGCGGCCCGGGGCTCCGGGAGAGAAAGCCGAGGGC	430	
Db	360	CGCGATGGCCCGGACGCGCCCGGACCGCGCGGCCCGGGGCTCCGGGAGAGAAAGCCGAGGGC	419	
Qy	431	GGGAGGCGGGGACTGCGCGGGACCTTCAGAGGAGAACCCGGGCGCGGAGAGAGCGGAGACC	490	
Db	420	GGGAGGCGGGGACTGCGCGGGACCTTCAGAGGAGAACCCGGGCGCGGAGAGAGCGGAGACC	479	
Qy	491	GCGGGGCGCCACCGGGGCTTGCCTGGGAGATGCTCGGTGCTTCCGATCCGCTTCAAGCGCC	550	
Db	480	GCGGGGCGCCACCGGGGCTTGCCTGGGAGATGCTCGGTGCTTCCGATCCGCTTCAAGCGCC	539	
Qy	551	AAGCGCTCCGAGAGCC-GGGTGCTTCGCCCTTCAGAGCACCTTCGCCCTTTCGACCGCGT	609	
Db	540	AAGCGCTCCGAGAGCCCGGGGTGCTTCGCCCTTCAGAGCACCTTCGCCCTTTCGACCGCGT	599	
Qy	610	GCTGGTG-ACGAGCAGAGGACATTACGACGCG-TCACGGGCAAGTTTCACTTCACAGTG	667	
Db	600	GCTGGTGAAACGAGCAGAGGACATTACGACGCGCTTTCACGGGCAAGTTTCACTTCACAGTG	659	
Qy	668	CTTGGGGCTTACTACTTCGCCGCTCATGCCACGGTCTACCGGGCGACGCTTCGAGTTTAT	727	
Db	660	CTTGGGGCTTACTACTTCGCCGCTCATGCCACGGTCTACCGGGCGACGCTTCGAGTTTAT	719	
Qy	728	CTGGTGAAGATGGCGAATCAATTGCTCTTTCTTCCAGTTTTCGGGGGGTGGCCCAAG	787	

Db	720	CTGTGTAAAGATAGCGCAATCATTTAGGTC--TTCTTCAGATTTTAAAGGGGGGTGGCC--AA	776
Qy	788	CCAGCCTCGCTCTCGGGGGGGGCCCATGTGTAGAGCTTGAGCCTTGAGACCAAGTGTGGGTG	847
Db	777	GCCAGCTCGCTCTCGGGGGGGGGCCATGTGTAGAGCTTGAGAGCTTGAGACCAAGTGTGGGTG	836
Qy	848	CAGGTGGGTC-TGGGTGACTCAATTTGGCATCTA	879
Db	837	CAGTGTGTGTGGGTGACTCACTGTGGCATCTA	869
RESULT 2			
LOCUS	BMS47549	1154 bp	mRNA linear EST 20-FEB-2002
DEFINITION	AGNCOURT 6507694 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5724522		
ACCESSION	BMS47549		
VERSION	BMS47549.1	GI:18781410	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1154)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Invitrogen cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNLML2713 row: 0 column: 19 High quality sequence stop: 591. Location/Qualifiers 1. 1154 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5724522" /clone_1Id="NIH_MGC_125" /lab_host="DH10B" /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6; Site 1: EcorV (destroyed); Site 2: NotI; RNA source pool of three ovaries, from females ranging in age from 38 to 49 yo. Library is oligo-dT primed and directionally cloned (EcorV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."		
FEATURES			
source			
BASE COUNT	178 a	382 c	401 g 193 t
ORIGIN			
Query Match	54.5%	Score 750.6;	DB 13; Length 1154;
Best Local Similarity	93.3%	Pred. No. 1.5e-93;	
Matches	818; Conservative	0; Mismatches	54; Indels 5; Gaps 3
Qy	40	CGGCAAGGAGCGAACCAAGACTGGGGGTGA	CGGCAAGGAGCGGCGGCGCTTGGCGGGGAGACA 99
Db	1	CGGCAAGGAGCGAACCAAGACTGGGGGTGA	CGGCAAGGAGCGGCGGCGGCGCTTGGCGGGGAGACA 60
Qy	100	ACCGGGGGGGCTTGAGCAACCAACTGAGGGGTCCGAGTAGTCGAGCGGCCCGGAAGAG	159
Db	61	ACCGGGGGGGCTTGAGCAACCAACTGAGGGGTCCGAGTAGTCGAGCGGCCCGGAAGAG	120
Qy	160	GCATTCGGGAGACCCGAGAGGGGGGACTGCGAAGGAGACCCCGGCGTTCGGGGCTTCCCGTGC	219
Db	121	GCATTCGGGAGACCCGAGAGGGGGGACTGCGAAGGAGACCCCGGCGTTCGGGGCTTCCCGTGC	180

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	B1490880/c										
139	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
187	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
199	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
247	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
259	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
307	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
319	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
367	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
379	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
427	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
439	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
487	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
499	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
547	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
558	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
607	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
618	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
666	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
678	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
726	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
737	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
786	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					
793	63031386771	NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3', mRNA sequence.	B1490880	B1490880	GI:15330108	human.					

Plate: L1A11430 row: f column: 17

High quality sequence start: 8
High quality sequence stop: 738.

FEATURES

source

1. 801
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5172880"
/clone_1ib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-Sport6; Site 1: NotI; site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT 180 a 251 c 237 g 133 t

ORIGIN

Query Match 51.8%; Score 712.6; DB 13; Length 801;
Best Local Similarity 98.2%; Pred. No. 2.7e-88;
Matches 773; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

527 CCTCCGCGATCCGCTTACG-CGCAAGCCTCCGAGAGCCGGGTGCTCCGCGCTGA 585
Db CCTTCGGATCCGCTTACGCGGCAAGCCTCCGAGAGCCGGGTGCTCCGCGCTGA 727
586 CGACCTTTCCTTTCGACCGCGCTGCTGTGAACGAG-CAGGACATTACGACCGCTCA 644
Db CGACCTTTCGCTTTCGACCGCGCTGCTGTGAACGAGCCGAGGACATTACGACCGCTCA 648
645 CCGGC-AGTTACCTGCGAGGAGGCTGCGGTCTACTACTTCCGCGGATGCGACCGCTC 703
Db CCGGCAAGTTACCTGCGAGGAGGCTGCGGTCTACTACTTCCGCGGATGCGACCGCTC 608
704 TACCGGCGACGCTGCGATTTGATCT-GGTGAAGATGGGAATTCATTCCTTTCTT 762
Db TACCGGCGACGCTGCGATTTGATCTGCGGAAGATGGGAATTCATTCCTTTCTT 548
763 CGAGTTTTCGGGGGGGGGCGGCAAGCCTGCTTCCGGGGGGGGGCAATGGTGAAGCT 822
Db CGAGTTTTCGGGGGGGGGCGGCAAGCCTGCTTCCGGGGGGGGGCAATGGTGAAGCT 488
823 GAGCCTGAGGACCAAGTGTGGGTGAGGTGGGTGAGTGACTACATTGGCATCTATGC 882
Db GAGCCTGAGGACCAAGTGTGGGTGAGGTGGGTGAGTGACTACATTGGCATCTATGC 428
883 GAGCATTACAGACAGACAGACCTTCTCCGAGATTCTGTGACTCCGACTGGACAGCTC 942
Db GAGCATTACAGACAGACAGACCTTCTCCGAGATTCTGTGACTCCGACTGGACAGCTC 368
943 CCAGCTTTCGTTAGTCCCACTGCAATGAGCTCAGTCTTCACTCTCTAAGAGAG 1002
Db CCAGCTTTCGTTAGTCCCACTGCAATGAGCTCAGTCTTCACTCTCTAAGAGAG 308
1003 GTGTGAGGCTGACACCAAGTGTATCCAGAGGCTGGCCCCCTGGAATTTGTGAATGA 1062
Db GTGTGAGGCTGACACCAAGTGTATCCAGAGGCTGGCCCCCTGGAATTTGTGAATGA 248
1063 CTAGGAGGTGGGGTGAAGACCTCTCCGCTCTGCTGTGAGCAAGAAATGGAAAGTGGC 1122
Db CTAGGAGGTGGGGTGAAGACCTCTCCGCTCTGCTGTGAGCAAGAAATGGAAAGTGGC 188
1123 TGCTGCGATCAGGTCTGCGAGCATGGGGGAGTGGCTGATTTCTGCGCAAGACAGAG 1182
Db TGCTGCGATCAGGTCTGCGAGCATGGGGGAGTGGCTGATTTCTGCGCAAGACAGAG 128
1183 AGTGTGCTGTGCGCAAGTGAATGCCCAAGTTGCTGTGCTCAAGAGCCCAAGGTGG 1242
Db AGTGTGCTGTGCGCAAGTGAATGCCCAAGTTGCTGTGCTCAAGAGCCCAAGGTGG 1242

Db 127 AGTGTGCTGTGCGCAAGTGAATGCCCAAGTTGCTGTGCTCAAGAGCCCAAGGTGG 68

Qy 1243 GGTGTCTCTCTTCCTGAGTCTCTGCTTCTGTGAATCTCCACCCCTCTGCTCTGG 1302

Db 67 GGTGTCTCTCTTCCTGAGTCTCTGCTTCTGTGAATCTCCACCCCTCTGCTCTGG 8

Qy 1303 GCGGCGC 1309

Db 7 GCGGCGC 1

RESULT 6 724 bp mRNA linear EST 25-SEP-2001

LOCUS B1771879 63055280F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204971 5'

DEFINITION mRNA sequence.

ACCESSION B1771879

VERSION B1771879.1 GI:15763457

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Butheria; Primates; Catarrhini; Homnidae; Homo.

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

plate: L1A11513 row: c column: 20

High quality sequence stop: 672.

Location/Qualifiers

1. 724

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5204971"

/clone_1ib="NIH_MGC_122"

/lab_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library"

BASE COUNT 176 a 184 c 211 g 152 t 1 others

ORIGIN

Query Match 48.8%; Score 672.2; DB 13; Length 724;
Best Local Similarity 99.4%; Pred. No. 9.2e-83;
Matches 674; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 700 CGTCTACCGGCGCAGCTGCGATTTGATCTGTGGAAGATGGGAATTCATTCCTTTT 759
Db 2 CGTCTACCGGCGCAGCTGCGATTTGATCTGTGGAAGATGGGAATTCATTCCTTTT 61
Qy 760 CTTCGAGTTTTCGGGGGGGGGCGGCAAGCAGGCTGCTCCGGGGGGGGGCAATGGTGA 819
Db 62 CTTCGAGTTTTCGGGGGGGGGCGGCAAGCAGGCTGCTCCGGGGGGGGGCAATGGTGA 121
Qy 820 GCTGAGCCTGAGGACCAAGTGTGGGTGAGGTGGGTGAGTGAATTCATTCCTTTT 879
Db 122 GCTGAGCCTGAGGACCAAGTGTGGGTGAGGTGGGTGAGTGAATTCATTCCTTTT 181

QY 880 TGCAGCATGACAGACAGACCTTCTCCGATTTCTGATGTAATCTCCAGCTGACAG 939
 DB 182 TGCAGCATGACAGACAGACCTTCTCCGATTTCTGATGTAATCTCCAGCTGACAG 241
 QY 940 CTCCTCCAGCTCTTCTGATGTAATCTCCAGCTGACAGCTGACAGCTGACAG 999
 DB 242 CTCCTCCAGCTCTTCTGATGTAATCTCCAGCTGACAGCTGACAGCTGACAG 301
 QY 1000 AGGCTGTGAGCTGACACCAAGCTGATCAGAGAGGCTGAGCTGAGATTTGTA 1059
 DB 302 AGGCTGTGAGCTGACACCAAGCTGATCAGAGAGGCTGAGCTGAGATTTGTA 361
 QY 1060 TGAAGTGAAG 1119
 DB 362 TGAAGTGAAG 421
 QY 1120 GAGCTGTGAGCTGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179
 DB 422 GAGCTGTGAGCTGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
 QY 1180 AGGATGTGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1239
 DB 482 AGGATGTGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
 QY 1240 TGGAGGTGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1299
 DB 542 TGGAGGTGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
 QY 1300 GGGGCGGCGCTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1359
 DB 602 GGGGCGGCGCTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
 QY 1360 AAAAAAAAAAAAAAAAAA 1377
 DB 662 AAAAAAAAAAAAAAAAAA 679

RESULT 7
 LOCUS BM920874 1045 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT_6706053 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752141
 ACCESSION BM920874
 VERSION BM920874.1 GI:19371253
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 plate: L1AM12785 row: n column: 14
 High quality sequence stop: 415.
 Location/Qualifiers
 1. 1045
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5752141"
 /clone_1kb="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:"

PCMV-SF016; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung; age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb.
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH MGC Library."

BASE COUNT 176 a 343 c 341 g 183 t 2 others
 ORIGIN

Query Match 48.7%; Score 671; DB 14; Length 1045;
 Best Local Similarity 95.4%; Pred. No. 1,1e-82;
 Matches 733; Conservative 0; Mismatches 31; Indels 4; Gaps 4;
 QY 171 GCCGAG 230
 DB 37 GCTGTGCTGACAG 96
 QY 231 GAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 290
 DB 97 GAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156
 QY 291 ACAAGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 350
 DB 157 ACAAGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216
 QY 351 GAGCAG 410
 DB 217 GAGCAG 276
 QY 411 CGGAG 470
 DB 277 CGGAG 336
 QY 471 CGGAG 530
 DB 337 CGGAG 396
 QY 531 CGGATCCGCTTCAAG 590
 DB 397 CGGATCCGCTTCAAG 456
 QY 591 CTTGCTCTTCAAG 650
 DB 457 CTTGCTCTTCAAG 516
 QY 651 AGTTCACTGCGAG 710
 DB 517 AGTTCACTGCGAG 576
 QY 711 CGAGCTGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 770
 DB 577 CGAGCTGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 636
 QY 771 TCGGAG 830
 DB 637 TCGGAG 696
 QY 831 AGGAG 887
 DB 697 ATGACCAAGT 756
 QY 888 TCAAGACAGACAGACCTTCT-CCGATTTCTGATGATGATGATGATGATGATGAT 934
 DB 757 TCAAGACAGACAGACCTTCTCCGATTTCTGATGATGATGATGATGATGATGAT 804

RESULT 8
 LOCUS BI770921 792 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603059737F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5209065 5',

mRNA sequence.
 ACCESSION B1770921
 VERSION B1770921.1 GI:15762499
 KEYWORDS EST.
 SOURCE Human.
 ORGANISM Homo sapiens
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>
 1 (bases 1 to 792)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Plate: L1AM1524 row: 3 column: 10
 High quality sequence stop: 782.
 Location/Qualifiers
 1..792
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5209065"
 /clone_1lb="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleen. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH MGC Library."
 BASE COUNT 142 a 232 c 237 g 180 t 1 others
 ORIGIN
 Query Match 48.0%; Score 660.4; DB 13; Length 792;
 Best Local Similarity 96.9%; Pred. No. 3.5e-81;
 Matches 758; Conservative 0; Mismatches 16; Indels 8; Gaps 8;
 QY 562 GAGCCGGGTGCTCCGCGCTCTGAGCCACCTTGGCTTCCAGCCGCGTGTGTGAACGA 621
 DB 1 GAGCCGGGTGCTCCGCGCTCTGAGCCACCTTGGCTTCCAGCCGCGTGTGTGAACGA 60
 QY 622 GAGGGGACATTACAGACGCGGTGACCGGGCAAGTTCACTGCGAGGTCCTGGGGGTCTACTA 681
 DB 61 GAGGGGACATTACAGACGCGGTGACCGGGCAAGTTCACTGCGAGGTCCTGGGGGTCTACTA 120
 QY 662 CTTCGCGCTGACATGCGACCGCTTACCGGGCCAGCTGAGTTGATCTGTGTAAGATAG 741
 DB 121 CTTCGCGCTGACATGCGACCGCTTACCGGGCCAGCTGAGTTGATCTGTGTAAGATAG 180
 QY 742 GCAATTCATGCTCTTCTTCCAGTTTTTGGGGGGGTGGCCCAAGCCAGCTGCTCTC 801
 DB 181 GCAATTCATGCTCTTCTTCTTCCAGTTTTTGGGGGGGTGGCCCAAGCCAGCTGCTCTC 240
 QY 802 GGGGGGGGGCCATGTGAGGCTGAGACCTGAGGACCAAGTGTGAGTGTGAGTGTGGG 861
 DB 241 GGGGGGGGGCCATGTGAGGCTGAGACCTGAGGACCAAGTGTGAGTGTGAGTGTGGG 300
 QY 862 TGACATCATTTGGCATCTATATGCGACATCAAGACAGACGACCTTCCGATTTCTGT 921
 DB 301 TGACATCATTTGGCATCTATATGCGACATCAAGACAGACGACCTTCCGATTTCTGT 360
 QY 922 GTACTCCGACTGGCAGAGCTCCGAGTCTTTGTAGTGGCCATGCAAGTGGAGCTCAT 981

DB 361 GTACTCCGACTGGCAGAGCTCCGAGTCTTTGTAGTGGCCATGCAAGTGGAGCTCAT 420
 QY 982 GCTCTCACTCTTACAGAGAGGGGTGTGAGGCTGACACCGATCATTCAGAGGGC-TGGC 1040
 DB 421 GCTCTCACTCTTACAGAGAGGGGTGTGAGGCTGACACCGATCATTCAGAGGGCCTGGC 480
 QY 1041 CCCCGTGAATATTTGTGAATATGATCTAGAGGAGGTGGGTGAGAC-ACCTCCGCTCTGCTGC 1099
 DB 481 CCCCGTGAATATTTGTGAATATGATCTAGAGGAGGTGGGTGAGACCACTTCCGCTCTGCTGC 540
 QY 1100 TGGCAGGAATGAGGACAGTGTGTCTGCGATCAGGCTGTGG-CAGCATGGGGAGTGGC 1158
 DB 541 TGGCAGGAATGAGGACAGTGTGTCTGCGATCAGGCTGTGGCCAGATGGGGAGTGGC 600
 QY 1159 TGAATTTCTGCGCCAGACACG-AGGATGTGTCTGTCTGCAAGTG-TAATGCCCACT 1216
 DB 601 TGAATTTCTGCGCCAGACACGAGGATGTGTCTGTGCGAAGTGTAACTCCCCCACT 660
 QY 1217 T-GGCTGTGTCAGAGGCCACG-GTGGGGGTCTCTCTT-CCTGGTCTCTGCTTCTGTG 1273
 DB 661 TAGCTCTGTGTCAGAGGCCACGAGTGAAGTCTTACCTGTGTCTGTCTGTCTGTG 720
 QY 1274 GATCCTCCCAACCCCTCTCTGCTCTCTGCGGGCCGCTTTTCTCAGAGATCATCATTA 1333
 DB 721 GATCCTCCCAACCCCTCTCTGCTCTCTGCGGGCCGCTTTTCTCAGAGATCATCATTA 780
 QY 1334 AC 1335
 DB 781 AC 782
 RESULT 9
 B1490062
 LOCUS B1490062
 DEFINITION 887 bp mRNA linear EST 28-AUG-2001
 603031867P1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5172880 5',
 mRNA sequence.
 ACCESSION B1490062
 VERSION B1490062.1 GI:15329290
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>
 1 (bases 1 to 887)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Plate: L1AM1430 row: 6 column: 17
 High quality sequence stop: 815.
 Location/Qualifiers
 1..887
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5172880"
 /clone_1lb="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C.


```

Db      241 CGGGGGCACTCCGGCTTCCAGGACGCGGGGCAACATGACGACGAGGCTTGCAGGCG 300
Qy      371 CGCAGTGGCCGCGACGCGCCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 430
Db      301 CGCAGTGGCCGCGACGCGCCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy      431 GGGAGGCGCGGACCTGCGAGGAGGACCCCGGGCGCGGACGAGAGGCGGAGGCGGAGGCGG 490
Db      361 GGGAGGCGCGGACCTGCGAGGAGGACCCCGGGCGCGGACGAGAGGCGGAGGCGGAGGCGG 420
Qy      421 GCGGGGCGCGGACCTGCGAGGAGGACCCCGGGCGCGGACGAGAGGCGGAGGCGGAGGCGG 550
Db      551 AGCGGCGCGGACCTGCGAGGAGGACCCCGGGCGCGGACGAGAGGCGGAGGCGGAGGCGG 610
Qy      481 AGCGGCGCGGACCTGCGAGGAGGACCCCGGGCGCGGACGAGAGGCGGAGGCGGAGGCGG 540
Db      611 CTGGTGAACGAGGAGGACCTTACGACGCGCGTCAACGCGGAGGAGGAGGAGGAGGAGGAG 670
Qy      541 CTGGTGAACGAGGAGGACCTTACGACGCGCGTCAACGCGGAGGAGGAGGAGGAGGAGGAG 600
Qy      671 GGGGCTCTACTACTTCCCGCTCCATGCGCACCG 701
Db      601 GGGGCTCTACTACTTCCCGCTCCATGCGCACCG 631

```

RESULT 11
 BG706609 630 bp mRNA linear EST 07-MAY-2001
 LOCUS 602672994F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795558 5'

DEFINITION
 mRNA sequence.
 BG706609
 VERSION BG706609.1 GI:13982121
 KEYWORDS
 EST.
 SOURCE
 ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 630)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLMI0678 row: 1 column: 23
 High quality sequence stop: 630.

FEATURES
 Source
 Location/Qualifiers

1. 630
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4795558"
 /clone_1b="NIH_MGC_96"
 /issue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTAA-3',
 size selected for average insert size 2.3 kb and
 normalized to R0F 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 93 a 215 c 257 g 65 t
 ORIGIN

Query Match 44.1%; Score 607.2; DB 12; Length 630;
 Best Local Similarity 99.5%; Pred. No. 7.1e-74;
 Matches 609; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy      7 TTCTCTTGAAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 66
Db      19 TCTCTTGAAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 78
Qy      67 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 126
Db      79 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 138
Qy      127 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 186
Db      139 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 198
Qy      187 GCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 246
Db      199 GCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 258
Qy      247 GCTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 306
Db      259 GCTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 318
Qy      307 CTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366
Db      319 CTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 378
Qy      367 GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
Db      379 GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 438
Qy      427 GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 486
Db      439 GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 498
Qy      487 ACCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 546
Db      499 ACCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 558
Qy      547 CGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606
Db      559 CGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 618
Qy      607 CGTGTGTGTGA 618
Db      619 CGTGTGTGTGA 630

```

RESULT 12
 BM893587/c 596 bp mRNA linear EST 29-APR-2002
 LOCUS 1126910.x1 Melton Normalized Human Islet 4 M4-HIS 1 Homo sapiens
 DEFINITION
 cDNA clone IMAGE:6136122 3' similar to TR:Q9UFK4 Q9UFK4
 HYPOHETICAL 22.8 KD PROTEIN ;, mRNA sequence.

ACCESSION
 BM893587
 VERSION
 BM893587.1 GI:19349055
 KEYWORDS
 EST.
 SOURCE
 ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 596)

REFERENCE
 AUTHORS
 Melton,D., Brown,D., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Scaerco,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Bliststein,A.,
 Schmitt,A., Theising,B., Ritzer,B., Ronko,I., Bennett,D., Cardenas
 M., Gibbons,M., McCann,R., Cole,R., Tsagarrelshvili,R., Williams,T.,
 Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium

JOURNAL
COMMENT

Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biolhp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
Seq primer: -40up from Gibco
High quality sequence stop: 441.

FEATURES
source

1. 596
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:6136122"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev stage="Adult"
/lab host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
Superscript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT 139 a 183 c 169 g 105 t
ORIGIN

Query Match 43.3%; Score 596; DB 14; Length 596;
Best Local Similarity 100.0%; Pred No. 2.5e-72;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

752 GCGTCTTTCTTCAGTTTTCGGGGGGGCGCCCAAGCCAGCTCGCTCGGGGGGGGCC 811
596 GCGTCTTTCTTCAGTTTTCGGGGGGGCGCCCAAGCCAGCTCGCTCGGGGGGGGCC 537
812 ATGTGAGGCTGAGAGCTTGAAGCAACAGTGTGGGTGAGGTGGGTGAGTACATTAATT 871
536 ATGTGAGGCTGAGAGCTTGAAGCAACAGTGTGGGTGAGGTGGGTGAGTACATTAATT 477
872 GGCATCTATGCGAGCATCAAGACAGACAGACCTTCTCCGATTTCTGGTACTCCGAC 931
476 GGCATCTATGCGAGCATCAAGACAGACAGACCTTCTCCGATTTCTGGTACTCCGAC 417
932 TGGCAAGCTCCCAAGTCTTTGCTTAGTCCCACTGCAAGTGAAGTCAATGCTCTACTC 991
416 TGGCAAGCTCCCAAGTCTTTGCTTAGTCCCACTGCAAGTGAAGTCAATGCTCTACTC 357
992 CTAGAAGAGGGGTGAGGCTGACAAACAGAGTCAACAGAGGCGTGGCCCCCTGGAAAT 1051
356 CTAGAAGAGGGGTGAGGCTGACAAACAGAGTCAACAGAGGCGTGGCCCCCTGGAAAT 297
1052 ATTGTGAATGACTAGGAGAGTGGGAGTGAAGCACTCTCCGTCTGCTGCTGCAAGAAATG 1111
296 ATTGTGAATGACTAGGAGAGTGGGAGTGAAGCACTCTCCGTCTGCTGCTGCAAGAAATG 237
1112 GGAACAGTGGCTGTCTGCGATCAAGTCTGGCAAGCATGGGGCAGTGGCTGAATTTCTGCC 1171
236 GGAACAGTGGCTGTCTGCGATCAAGTCTGGCAAGCATGGGGCAGTGGCTGAATTTCTGCC 177

QY 1172 AAGACACAGAGAGTGTCTGTGCTGAGCAAGTGAATGCTCCCAAGTTGCTGTGTCAGAGA 1231
DB 176 AAGACACAGAGAGTGTCTGTGCTGAGCAAGTGAATGCTCCCAAGTTGCTGTGTCAGAGA 117
QY 1232 GCCACAGGTGGGGGTCTCTCTTCCTGGGTCCTGTGCTTCTGGAATCCCTCCCAAGCCCTTC 1291
DB 116 GCCACAGGTGGGGGTCTCTCTTCCTGGGTCCTGTGCTTCTGGAATCCCTCCCAAGCCCTTC 57

QY 1292 CTGCTCTCGGGGCGCGCCCTTTTCTCAGAGATCACTCAATAAAGCTTAAGAAACCTTC 1347
DB 56 CTGCTCTCGGGGCGCGCCCTTTTCTCAGAGATCACTCAATAAAGCTTAAGAAACCTTC 1

RESULT 13
BM544255 596 bp mRNA linear EST 20-FEB-2002
LOCUS
DEFINITION AGENCOURT 6490655 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587759
ACCESSION
VERSION BM544255.1 GI:18775356
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM12357 row: m column: 08

High quality sequence stop: 595.
Location/Qualifiers

FEATURES
source

1. 596
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5587759"
/clone_lib="NIH MGC_125"
/lab host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1.3-5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 138 a 158 c 172 g 127 t 1 others
ORIGIN

Query Match 42.6%; Score 586.4; DB 13; Length 596;
Best Local Similarity 99.7%; Pred No. 5e-71;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

789 CAGCTCGCTCTCGGGGGGGGCGCATGTGTGAGGCTGAGAGCTGAGAGCAAGTGTGGTGC 848
DB 1 CAGCTCGCTCTCGGGGGGGGCGCATGTGTGAGGCTGAGAGCTGAGAGCAAGTGTGGTGC 60
QY 849 AGTGGGTGTGGGTGATCACTTAATTTGCAATTCATGCAAGATCAAGACAGACAGACCTTCT 908
DB 61 AGTGGGTGTGGGTGATCACTTAATTTGCAATTCATGCAAGATCAAGACAGACAGACCTTCT 120
QY 909 CCGGATTTCTGTGTACTCCGACTGCGACAGCTCCCAAGTCTTTGCTTAGTGGCCAGTGC 968

```

Db      121 CCGAGTTCTGTGTAATCTCCAGCTGCGACAGCTCCCACTTTGTTAGTCCCACTGC 180
Qy      969 AAGTAGAGTCAATGCTCTCACTCTAGAGAGAGGTGTAGAGCTCAACAGAGTATCC 1028
Db      181 AAGTAGAGTCAATGCTCTCACTCTAGAGAGAGGTGTAGAGCTCAACAGAGTATCC 240
Qy      1029 AAGAGAGGCTGGCCCCCTGGAAATATTGTGAAATAGTAAAGAGAGTGGGTAGAGCACTTC 1088
Db      241 AGAAGGGCTGGCCCCCTGGAAATATTGTGAAATAGTAAAGAGAGTGGGTAGAGCACTTC 300
Qy      1489 CTTCTGCTGCTGCGAAGAGAAATGGAACAGTGTCTGTGCGATCAGTCTGGCAGCATG 1148
Db      301 CTTCTGCTGCTGCGAAGAGAAATGGAACAGTGTCTGTGCGATCAGTCTGGCAGCATG 360
Qy      1149 GGGCAGTGTGCTGATTTCTGCCCCAGACAGAGAGTGTGTGTCTGTGCGAAGTATAGT 1208
Db      361 GGGCAGTGTGCTGATTTCTGCCCCAGACAGAGAGTGTGTGTCTGTGCGAAGTATAGT 420
Qy      1209 CCCCCAGTGTCTGTGTCCAGAGAGCCAGGTGGGTGTCTGTCTGTCTGTCTGTCTGTCT 1268
Db      421 CCCCCAGTGTCTGTGTCCAGAGAGCCAGGTGGGTGTCTGTCTGTCTGTCTGTCTGTCT 480
Qy      1269 CTTGTGATCTCTCCCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1328
Db      481 CTTGTGATCTCTCCCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Qy      1339 AATTAACCTTAGAACCTCTATATAAAAAAAAAAAAAAAAAAAAAA 1377
Db      541 AATTAACCTTAGAACCTCTATATAAAAAAAAAAAAAAAAAAAAAA 589

```

```

RESULT 14
LOCUS      Bi821899          941 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603035796F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176967 5',
ACCESSION  Bi821899
VERSION     Bi821899.1 GI:15933449
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 941)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cga@pds-rcmail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/BLNL at:
            http://image.llnl.gov
            Plate: LHAM11440 row: p column: 24
            High quality sequence stop: 779.
            Location/Qualifiers
                1. 941

```

FEATURES

```

source
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5176967"
    /clone_1ib="NIH_MGC_115"
    /lab_host="DH10B"
    /note="Organ: pooled brain, lung, testis; Vector:
    pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
    source anonymous pool of 6 male brains, age range 23-27; 1
    male lung, age 27; and 1 male testis, age 69. Library is
    oligo-dT primed and directionally cloned (EcoRV site is
    destroyed upon cloning). Average insert size 1.8 kb,
    insert size range 1-3 kb. Library is normalized and

```

enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC library."

BASE COUNT

130 a 320 c 329 g 162 t

ORIGIN

Query Match 42.4%; Score 583.2; DB 13; Length 941;

Best Local Similarity 94.3%; Pred. No. 1e-70; 28; Indels 18; Gaps 14;

Matches 758; Conservative 0; Mismatches 28; Indels 18; Gaps 14;

```

Qy      191 GAGAACCCCGGCTCCGAGCTCCGAGTCCAGAGCTATAG- GCCACTCTCTCTCTCTCT 249
Db      136 GAGAACCCCGGCTCCGAGCTCCGAGTCCAGAGCTATAGAGCACTCTCTCTCTCTCT 195
Qy      250 GCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309
Db      196 GCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 255
Qy      310 CCGGAGGACACCCGAGCTTCCAGAGACCGAGGACCAACATGAGACCA- GGGCTTGGCCG 368
Db      256 CCGGAGGACACCCGAGCTTCCAGAGACCGAGGACCAACATGAGACCACTGGCTTGGCCG 315
Qy      369 GCGCGATGAGCGCGACAGCGCGCGCGCG- GGGCTTCCGAGAGAGAGCGAG 427
Db      316 GCGCGATGAGCGCGACAGCGCGCGCGCG- GGGCTTCCGAGAGAGAGCGAG 375
Qy      428 GCGCGAGAGCGCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487
Db      376 GCGCGAGAGCGCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 435
Qy      488 CCGGAGGAGCGCGAGCGAG- CTTGCGGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCT 545
Db      436 CCGGAGGAGCGCGAGCGAG- CTTGCGGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCT 495
Qy      546 GCGCGAGAGCGCTTCCAGAGAC- GGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 602
Db      496 GCGCGAGAGCGCTTCCAGAGAC- GGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 555
Qy      603 ACCGCGTGTCTGTGAACGAGAGGAGCACTTACAGCGCGGTGACCGGCAAGTCACTGTGC 662
Db      556 ACCGCGTGTCTGTGAACGAGAGGAGCACTTACAGCGCGGTGACCGGCAAGTCACTGTGC 615
Qy      663 -AGTGTCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Db      616 GAGTGTCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 675
Qy      721 GTTGTGTGTGTGTGAAGATGCGAATTCATTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db      676 GTTGTGTGTGTGTGAAGATGCGAATTCATTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 734
Qy      781 GCGCAAGCAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 839
Db      735 GCGCAAGCAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 794
Qy      840 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 898
Db      795 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 853
Qy      899 AGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 958
Db      854 AGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 912
Qy      959 TGCCCACTGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 912
Db      913 ---CCCTGTCAAGTAGTCTCTCTG 933

```

```

RESULT 15
LOCUS      BM893691/c          578 bp      mRNA      linear      EST 29-APR-2002
DEFINITION j128e08.x1 Melton Normalized Human Retlet 4 NA-HIS 1 Homo sapiens
            cDNA clone IMAGE:6136046 3', similar to TR:Q9UFK4 Q9UFK4
            HYPOTHETICAL 22.8 KD PROTEIN ;, mRNA sequence.

```

ACCESSION	BM893691	GI:19349159
VERSION	BM893691.1	
KEYWORDS	EST.	
SOURCE	human.	

FEATURES
source

BASE COUNT	130 a	181 c	163 g	104 t
ORIGIN				

Query Match	42.0%;	Score 578;	DB 14;	length 578;
Best Local Similarity	100.0%;	Pred. No. 7.1e-70;		
Matches 578;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

830 GAGGACCAAGTGTGGTGCAGGTGGGTGTGGTACTAATTGGCATCTATGCCAGCATC 889
 518 GAGACCAAGTGTGGTGCAGGTGGGTGTGGTACTAATTGGCATCTATGCCAGCATC 459

0y 830 AAGCAGACAGCAGCTTCGCGATTCGCTGATCTCGCATGGCAGAGTCCCAAGTC 949
 Db 458 AAGCAGACAGCAGCTTCGCGATTCGCTGATCTCGCATGGCAGAGTCCCAAGTC 399

Qy	950	TTTGCTTAGTGCCCACTGCAAAAGGAGCTCATGCTCTCATCTCCAGAAAGAGGGGTGGAG	1009
Db	398	TTTGTCTTAGTGCCCACTGCAAAAGGAGCTCATGCTCTCATCTCCAGAAAGAGGGGTGGAG	339
Qy	1010	GCTGACAAACCAAGGTCATCCAGAGAGGCTGGCCCCCTCGAATATTGTGAATGACTAGAGGA	1069
Db	338	GCTGACAAACCAAGGTCATCCAGAGAGGCTGGCCCCCTCGAATATTGTGAATGACTAGAGGA	279
Qy	1070	GGTGGGGGTAGAGCACTTCCGTCCTCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGC	1129
Dc	278	GGTGGGGGTAGAGCACTTCCGTCCTCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGC	219
Qy	1130	GATCAGGCTGGGCGAGCATGGGGGCGAGTGGCGGAATTTCTGCCCCAAGACCAAGAGAGTGGC	1189
Db	218	GATCAGGCTGGGCGAGCATGGGGGCGAGTGGCGGAATTTCTGCCCCAAGACCAAGAGAGTGGC	159
Qy	1190	TGTGCTGCGCAAGTGAATGATCCCCAGTGTGCTTGGTTCAGAGAGCCCAAGTGGGGGTCTC	1249
Db	158	TGTGCTGCGCAAGTGAATGATCCCCAGTGTGCTTGGTTCAGAGAGCCCAAGTGGGGGTCTC	99
Qy	1250	TCTTCTGCTGTCCTCTGCTTCTCTGGAATCTTCCCCACCCCTCTGCTCTGGGGGCTGGCC	1309
Db	98	TCTTCTGCTGTCCTCTGCTTCTCTGGAATCTTCCCCACCCCTCTGCTCTGGGGGCTGGCC	39
Qy	1310	CTTTTCTCAGAGATCACTCAATAAACCCTAAGAACCCCTC	1347
Db	38	CTTTTCTCAGAGATCACTCAATAAACCCTAAGAACCCCTC	1

Search completed: June 20, 2003, 08:43:41
Job time : 1908 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:26:08 ; Search time 36 Seconds
(without alignments)
899.442 Million cell updates/sec

Title: US-09-944-403-42

Perfect score: 243
Sequence: 1 MRPLVLLTGLAGSPPLD.....DSTFGSLVYSDWHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002:*

1	_	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1980.DAT *
2	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1981.DAT *	
3	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1982.DAT *	
4	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1983.DAT *	
5	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1984.DAT *	
6	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1985.DAT *	
7	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1986.DAT *	
8	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1987.DAT *	
9	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1988.DAT *	
10	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1989.DAT *	
11	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1990.DAT *	
12	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1991.DAT *	
13	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1992.DAT *	
14	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1993.DAT *	
15	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1994.DAT *	
16	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1995.DAT *	
17	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1996.DAT *	
18	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1997.DAT *	
19	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1998.DAT *	
20	/SIDS2/gcgdata/genseq/genseqp-emb1/AA1999.DAT *	
21	/SIDS2/gcgdata/genseq/genseqp-emb1/AA2000.DAT *	
22	/SIDS2/gcgdata/genseq/genseqp-emb1/AA2001.DAT *	
23	/SIDS2/gcgdata/genseq/genseqp-emb1/AA2002.DAT *	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	243	100.0	243	20	AAV06481	Human tumour-associ
2	243	100.0	243	20	AAV17827	Human PRO344 prote
3	243	100.0	243	20	AAW97984	Human adipocyte-sp
4	243	100.0	243	21	AAH33461	Human PRO344 prote
5	243	100.0	243	21	AAV71468	Human PRO344 prote
6	243	100.0	243	21	AAV93688	Amino acid sequenc
7	243	100.0	243	21	AAH01318	Human PRO344 polyV
8	243	100.0	243	22	AAU12352	Human PRO344 polyP
9	243	100.0	243	22	AAH65815	Human TRANCO 253 SE
10	243	100.0	243	22	AAH49593	Human adipocyte co

1	243	100.0	243	22	AAB84599	Human adipocyte co
12	228	93.8	228	22	AAB65816	Human mature TANGO
13	213	87.7	220	22	ABG12724	Novel human diagno
14	185	76.1	243	22	AAB65888	Human secreted pro
15	174	71.6	243	22	AAB65891	Human secreted pro
16	173	70.2	201	23	AAO21653	Human secreted pro
17	148	60.9	243	22	AAB65889	Human secreted pro
18	142	58.4	243	22	AAB65890	Human secreted pro
19	128	52.7	128	22	AAB65819	Human TANGO 253 C1
20	82	33.7	151	22	ABG12723	Novel human diagno
21	70	28.8	243	21	AAI76040	Rat skin cell proteo
22	70	28.8	243	22	AAB55979	Skin cell protein, R
23	70	28.8	243	23	ABR72179	Rat protein isolat
24	62	25.5	228	22	AAB65821	Murine mature TANGO
25	62	25.5	242	22	AAB65896	Murine secreted pr
26	62	25.5	243	22	AAB65890	Murine TANGO 253 S
27	62	25.5	243	22	AAB65897	Murine secreted pr
28	62	25.5	243	22	AAB65898	Murine secreted pr
29	62	25.5	243	22	AAB65899	Murine secreted pr
30	60	24.7	60	22	AAB65814	Human TANGO 253 co
31	51	21.0	128	22	AAB65828	Murine TANGO 253 C
32	43	17.7	60	22	AAB65823	Murine TANGO 253 C
33	33	12.3	87	21	AAI87269	Human signal peptid
34	15	6.2	15	22	AAB65817	Human TANGO 253 s1
35	11	4.5	27	22	AAB92425	Miscellaneous pept
36	11	4.5	56	22	ABR11339	Peptide #1390 encco
37	11	4.5	56	22	ABR36545	Peptide #4051 encco
38	11	4.5	56	22	ABR21884	Protein #1383 encco
39	11	4.5	56	22	AAI57307	Human brain expres
40	11	4.5	56	22	AAI69711	Human bone marrow
41	11	4.5	56	22	AAI17524	Peptide #1398 encco
42	11	4.5	56	22	AAI30045	Peptide #4082 encco
43	11	4.5	56	22	AAI05136	Peptide #1378 encco
44	11	4.5	56	23	ABG39359	Human secreted prote
45	11	4.5	128	20	AAI02650	Human secreted prote

ALIGNMENTS

RESULT 1
ID AAY06481
AAY06481 standard; Protein; 243 AA.
XX AC AAY06481;
XX DT 27-SEP-1999 (first entry)
XX DE Human tumour-associated protein PRO344.
XX KM PRO344; UNQ303; cancer; tumour; diagnosis; therapy; human
XX OS Homo sapiens.
XX Key
FH Peptide
FT Location/Qualifiers
FT 1..15
FT /note= "signal peptide"
FT 16..243
FT /note= "mature protein"
FT 68..215
FT /note= "N-myristoylated"
FT 216..243
FT /note= "N-myristoylated"
XX Modified-site
XX EN *W09935170-A2.
XX PD 15-JUL-1999.
XX PF 05-JAN-1999;
XX PF 99WO-US00106.
XX PR 20-NOV-1998;
PR 98US-0109304.
PR 05-JAN-1998;
PR 98US-0070440.
PR 29-APR-1998;
PR 98US-0083500.

PR 22-MAY-1998; 98US-0086414.
 PR 10-JUN-1998; 98US-0086742.
 PR 10-NOV-1998; 98US-0107783.
 XX
 PA (GETH) GENENTECH INC.
 PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA,
 PI Roy M, Wood WI,
 DR WPI; 1999-430385/36.
 DR N-PSDB; AAX87258.
 XX
 PT Antibody against proteins expressed in neoplastic cells, useful for
 PT tumor diagnosis and treatment
 PS
 PS Example 1; Fig 10; 162pp; English.
 XX
 CC This sequence represents human PRO344 (UN0303), a protein encoded
 CC by the novel cDNA clone DNA40592 (see AAX87258). Amplification of
 CC DNA40592 was observed in primary lung tumours and in primary colon
 CC tumours, suggesting a significant role in tumour formation and
 CC growth. Antagonists (e.g. antibodies) directed to PRO344 may have
 CC use in cancer therapy. The invention identifies 14 genes (see
 CC AAX87254-67) that are amplified in the genome of tumour cells. Such
 CC amplification is expected to be associated with overexpression of
 CC the gene product and to contribute to tumorigenesis. The encoded
 CC proteins (see AAY06477-90) may be useful targets for the diagnosis
 CC and/or treatment (including prevention) of certain cancers, and may
 CC act as predictors of the prognosis of tumour treatment. Antibodies
 CC that bind the proteins are claimed and used in claimed cancer
 CC diagnostic kits.
 XX
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 243; DB 20; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHHSQGLPGDGRDGRDAPG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHHSQGLPGDGRDGRDAPG 60
 QY 61 APGEKGGGRPGPLPGRGDPGPRGEGAPGPTGAGECSVPPRSASAKRSRRVPPSD 120
 DB 61 APGEKGGGRPGPLPGRGDPGPRGEGAPGPTGAGECSVPPRSASAKRSRRVPPSD 120
 QY 121 APLPFDRLVNEQGHDAVTGKFTCOVPGVYFAVAATYRASLQFDLVNKGESIASFPQ 180
 DB 121 APLPFDRLVNEQGHDAVTGKFTCOVPGVYFAVAATYRASLQFDLVNKGESIASFPQ 180
 QY 181 FFGGMPKPASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDWHSRP 240
 DB 181 FFGGMPKPASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDWHSRP 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 2
 AAY17827
 ID AAY17827 standard; Protein; 243 AA.
 XX
 XX AAY17827;
 AC
 XX
 DT 12-AUG-1999 (first entry)
 XX
 XX Human PRO344 protein sequence.
 DE
 XX
 KM Human; PRO protein; tumour necrosis factor family; TNF; cytokine;
 XX secreted protein; transmembrane protein; inflammation disorder.
 OS Homo sapiens.

XX
 FN W09928462-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 01-DEC-1998; 98MO-US25108.
 XX
 PR 25-FEB-1998; 98US-0075945.
 PR 03-DEC-1997; 97US-0067411.
 PR 11-DEC-1997; 97US-0069278.
 PR 11-DEC-1997; 97US-0069334.
 PR 11-DEC-1997; 97US-0069335.
 PR 12-DEC-1997; 97US-0069425.
 PR 16-DEC-1997; 97US-0069694.
 PR 16-DEC-1997; 97US-0069696.
 PR 16-DEC-1997; 97US-0069702.
 PR 17-DEC-1997; 97US-0069870.
 PR 17-DEC-1997; 97US-0069873.
 PR 18-DEC-1997; 97US-0068017.
 PR 05-JAN-1998; 98US-0070440.
 PR 09-FEB-1998; 98US-0074086.
 PR 09-FEB-1998; 98US-0074092.
 XX
 PA (GETH) GENENTECH INC.
 PI Baker KP, Chen J, Goddard A, Gurney AL, Wood WI,
 PI Yuan J;
 XX
 DR WPI; 1999-371118/31.
 DR N-PSDB; AAX80052.
 XX
 FT Nucleic acids encoding PRO secreted and transmembrane proteins
 PS
 PS Claim 12; Fig 21; 123pp; English.
 CC The present invention describes nucleic acids encoding PRO secreted and
 CC transmembrane proteins used therapeutically. The PRO proteins have
 CC cytostatic, anti-inflammatory, anti-proliferative and immunosuppressive
 CC activity. The proteins and polynucleotides can be used in therapy.
 CC identification of homologues, raising antibodies and design of probes
 CC and primers. They can be used in a range of diseases related to proteins
 CC that they have homology with, e.g. a PRO protein having homology to
 CC complement proteins may be used in inflammatory responses.
 XX
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 243; DB 20; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHHSQGLPGDGRDGRDAPG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHHSQGLPGDGRDGRDAPG 60
 QY 61 APGEKGGGRPGPLPGRGDPGPRGEGAPGPTGAGECSVPPRSASAKRSRRVPPSD 120
 DB 61 APGEKGGGRPGPLPGRGDPGPRGEGAPGPTGAGECSVPPRSASAKRSRRVPPSD 120
 QY 121 APLPFDRLVNEQGHDAVTGKFTCOVPGVYFAVAATYRASLQFDLVNKGESIASFPQ 180
 DB 121 APLPFDRLVNEQGHDAVTGKFTCOVPGVYFAVAATYRASLQFDLVNKGESIASFPQ 180
 QY 181 FFGGMPKPASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDWHSRP 240
 DB 181 FFGGMPKPASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDWHSRP 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 3
 AAM97984

	CC	methionine start site, a hydrophobic region of approximately 13 amino acids and a cleavage site. A single EST sequence was discovered, and the novel polypeptide encoded by the full-length cDNA allowed the identification of a homologue relationship with adipocyte complement related protein Acrp30 and adipocyte secreted protein apM1. A full-length clone (see AAZ4684) was obtained from a lung tissue library. Expression vectors, cultured cells and a method of producing zsig99 polypeptides are claimed, as well as zsgis99 polypeptides having N- or C-terminal affinity tags, toxins, radionucleotides, enzymes or fluorophores, fusion proteins including zsgis99 polypeptides, an antibody that specifically binds to an epitope of zsgis99, and a method for modulating free fatty acid metabolism by administering a zsgis99 polypeptide.
	CC	The zsgis99 polypeptide may also be used in organ preservation, for cryopreservation, for surgical pretreatment to prevent injury due to ischaemia and/or inflammation, and as an antimicrobial agent, promoting lysis or phagocytosis of infectious agents.
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
	CC	
SQ	Sequence	243 AA;
	Query Match	100.0%; Score 243; DB 20; Length 243;
	Best Local Similarity	100.0%; Pred. No. 4,9e-212;
	Matches:	243; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY		1 MRPLLVLTLTGLAAGSPPLDNDKIPSLCRGHGPGTPFGHHSGOCLPDRGRDGHDGAPG DB 1 MRPLLVLTLTGAAASPPDLNKKIPLSCLPHGPLGTFTGHHSOGCLPDRGRDGRDGA PG 60
OY		61 APGEKGEGRGRLPEPRDDPGPRGEPAGEPTGPAGECSVPFRSAFSAKRSESRRVPPSD DB 61 APGEKGEGRGRLPEPRDDPGPRGEPAGEPTGPAGECSVPFRSAFSAKRSESRVPPSD 120
OY		121 APLPEDRVLVNEGGHYDAVTGKTCTGVPCGYTFEVAHATVVTAASLDFLVNGESIASFPQ DB 121 APLPEDRVLVNEGGHYDAVTGKTCTGVPCGYTFEVAHATVVTAASLDFLVNGESIASFPQ 180
OY		181 FFCGGPKRPAISGGGMWVRLEPDWDVVGWGVDYGIYAISKIDSTFSGPSLYPSDMHSSP DB 181 FFCGGPKRPAISGGGMWVRLEPDWDVVGWGVDYGIYAISKIDSTFSGPSLYPSDMHSSP 240
OY		241 VFA 243 DB 241 VFA 243
	RESULT 4	AAB33461
ID	AAB33461 standard; Protein; 243 AA.	
XX	AAB33461;	
DT	29-JAN-2001 (First entry)	
DE	Human PRO344 protein UNQ303 SEQ ID NO:241.	
KX	Human; immune related disease; diagnosis; antiinflammatory; cardiac; dermatological; antithyroidic; antirheumatic; immunosuppressive; haemostatic; antihypoid; antidiabetic; nocrotropic; neuroprotective; antianaemic; hepatotropic; vincicide; anticpsoriatic; anti allergic; antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis; spondylarthropathy; systemic sclerosis; sarcoidosis; idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis; systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus; autoimmune thrombocytopoenia; immune-mediated renal disease; demyelinating bowel disease; hepatobiliary disease; Whipple's disease; inflammatory bowel disease; gluten-sensitive enteropathy; autoimmune disease; immune-mediated skin disease; allergy disease; immunologicall disease; transplantation associated disease; graft rejection; graft-versus-host-disease.	
OS	Homo sapiens.	
PJ	WO20005758-A2.	

XX		14-SEP-2000.	
PD			
XX			
PF	02-MAR-2000;	2000MO-US05841.	
XX			
PR	08-MAR-1999;	99WO-US05028.	
XX	10-MAR-1999;	99US-0123618.	
PR	12-MAR-1999;	99US-0123957.	
PR	23-MAR-1999;	99US-0125775.	
PR	12-APR-1999;	99US-0128849.	
PR	20-APR-1999;	99WO-US08615.	
PR	28-APR-1999;	99US-0131445.	
PR	04-MAY-1999;	99US-0132371.	
PR	14-MAY-1999;	99US-0134287.	
PR	02-JUN-1999;	99WO-US12252.	
PR	23-JUN-1999;	99US-0141037.	
PR	20-JUL-1999;	99US-0144756.	
PR	26-JUL-1999;	99US-0145698.	
PR	28-JUL-1999;	99US-0146222.	
PR	01-SEP-1999;	99WO-US20111.	
PR	08-SEP-1999;	99WO-US20594.	
PR	13-SEP-1999;	99WO-US20944.	
PR	15-SEP-1999;	99WO-US21099.	
PR	15-SEP-1999;	99WO-US21547.	
PR	05-OCT-1999;	99WO-US23089.	
PR	29-OCT-1999;	99US-0162506.	
PR	29-NOV-1999;	99WO-US28214.	
PR	30-NOV-1999;	99WO-US28313.	
PR	01-DEC-1999;	99WO-US28409.	
PR	01-DEC-1999;	99WO-US28301.	
PR	01-DEC-1999;	99WO-US28634.	
PR	02-DEC-1999;	99WO-US28551.	
PR	02-DEC-1999;	99WO-US28564.	
PR	16-DEC-1999;	99WO-US28565.	
PR	16-DEC-1999;	99WO-US30095.	
PR	20-DEC-1999;	99WO-US30999.	
PR	30-DEC-1999;	99WO-US31274.	
PR	05-JAN-2000;	2000MO-US00213.	
PR	06-JAN-2000;	2000MO-US00277.	
PR	06-JAN-2000;	2000MO-US00376.	
PR	11-FEB-2000;	2000MO-US03565.	
PR	18-FEB-2000;	2000MO-US04341.	
PR	18-FEB-2000;	2000MO-US04342.	
PR	22-FEB-2000;	2000MO-US04414.	
XX			
PA	(GENTH) GENEINTECH INC.		
PI	Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W,		
P1	Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V,		
P1	Stewart TJ, Tunas D, Watanabe CK, Wood WI, Yan M,		
DR	WPI, 2000-572271/53.		
DR	N-PSDB; AAC58626.		
XX			
PT	Sixty four PRO polypeptides, useful in the diagnosis and treatment of		
PT	sixty related disorders, e.g. systemic lupus erythematosus, rheumatoid		
PT	arthritis, osteoarthritis, thyroiditis and diabetes mellitus -		
XX			
PS	Claim 33; Fig 96; 309pp; English.		
CC	The present invention describes sixty four human PRO proteins which can		
CC	be used in the treatment of immune related diseases. The human PRO		
CC	proteins, anti-PRO antibodies, agonists and antagonists are useful for		
CC	treating and diagnosing immune related disorders. The disorders are		
CC	selected from systemic lupus erythematosus, rheumatoid arthritis,		
CC	osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,		
CC	systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's		
CC	syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic		
CC	anemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,		
CC	immune-mediated renal disease, demyelinating diseases of the central		
CC	and peripheral nervous systems, hepatobiliary diseases, inflammatory		
CC	bowel disease, gluten-sensitive enteropathy and Whipple's disease,		
CC	autoimmune or immune-mediated skin diseases, allergic diseases,		

	CC	immunological diseases of the lung, and transplantation associated diseases including graft rejection and graft-versus-host-disease.
	CC	AAC5897 to AAC5878 represent PCR primers and hybridisation probes used in the isolation of human PRO sequences. AAC5879 to AAC58642 and
	CC	AAB3414 to AAB3477 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention.
XX	SQ	Sequence 243 AA;
	Query Match	100.0%; Score 243; DB 21; Length 243;
	Best Local Similarity	100.0%; Pred. No. 4.9e-212; Mismatches 0; Gaps 0
	Matches 243;	Conservative 0; Indels 0;
Dy	1	MRPLVLLGLLAAGSPPLDDNKLPSLCPGHPLGTPGHHSOGILPGRDGRDGAPG 60
Dy	1	MRPLVLLGLLAAGSPPLDDNKLPSLCPGHPLGTPGHHSOGILPGRDGRDGAPG 60
Oy	61	APGEKGGGRPGLPFGPRGDPCGRBAGPAGTGPAGECSVPPRSAPSAKRSSRPVPPSD 120
Dy	61	APGEKGGGRPGLPFGPRGDPCGRBAGPAGTGPAGECSVPPRSAPSAKRSSRPVPPSD 120
Oy	121	APLPEPDVLVNEOQHVAIVNGKFICOVPGVVYFPVHAATVYSALQFDLVKNESTIASFPQ 180
Dy	121	APLPEPDVLVNEOQHVAIVNGKFICOVPGVVYFAVHAATVYSALQFDLVKNESTIASFPQ 180
Oy	181	FFGGMPKPASISGGAMVRLBEDDVQWVGVDYIGIYASIKTDSTFSGLVYSDMHSSP 240
Dy	181	FFGGMPKPASISGGAMVRLBEDDVQWVGVDYIGIYASIKTDSTFSGLVYSDMHSSP 240
Oy	241	VFA 243
Dy	241	VFA 243
	RESULT 5	
ID	AA71468	standard; Protein; 243 AA.
XX	AA71468;	
AC	08-NOV-2000	(first entry)
DT		
XX		
DB	Human PRO344 protein.	
XX	PRO344; DNA0592-1242; human; ATCC No: 209492; antiproliferative;	
KM	neoplastic cell growth inhibitor; cytoskeletal; treatment; cancer; tumour;	
KM	breast; prostate; colon; lung; renal; ovarian; central nervous system;	
KM	CNS; leukemia; melanoma; Expressed Sequence Tag; EST; secreted protein;	
KM	extracellular domain; ECD.	
XX	Homo sapiens.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..15
FT	Modified-site	/label= Signal_peptide
FT	Protein	11..17
FT	Modified-site	/note= "N-myristoylation site"
FT	Modified-site	16..243
FT	Binding-site	/label= Mature_PRO344_protein
FT	Modified-site	68..74
FT	Modified-site	/note= "N-myristoylation site"
FT	Modified-site	77..80
FT	Modified-site	/note= "Cell attachment .sequence"
FT	Modified-site	216..222
FT	Modified-site	/note= "N-myristoylation site"
XX	WO200032778-A2.	
XX	08-JUN-2000.	
XX	30-NOV-1999;	99WO-US28409.
PF	01-DEC-1998;	98WO-US25108.
XX		

PR 16-DEC-1998; 98US-0112850.
 PR 22-DEC-1998; 98US-0113296.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PA (GENTH) GENENTECH INC.
 PI Chen J, Goddard A, Gurney AL, Hillan K, Napier M, Wood WI;
 XX WPI; 2000-412325/35.
 DR N-PSDB; AAD01241.
 XX
 XX New composition useful for inhibiting neoplastic cell growth and for
 PT treating cancers, comprises PRO655, PRO344 or PRO364 polypeptide or
 PT their antagonists
 XX
 XX Claim 31; Fig 6; 108pp; English.
 PS
 XX The present sequence is the human PRO344 protein, encoded by the cDNA
 CC clone, designated as DNA40592-1242. It is isolated from human foetal
 CC lung tissue, cDNA library, identified using probes based on a consensus
 CC sequence DNA4398, derived from secreted protein extracellular domain
 CC (BCD) expressed sequence tag (EST). This clone is assigned ATCC deposit
 CC No: 209492. PRO344 functions as a neoplastic cell growth inhibitor and
 CC is used for treating tumours, using an effective amount of PRO655, PRO364
 CC and PRO344. This composition is especially useful for treatment of human
 CC cancers such as breast, prostate, colon, lung, renal, ovarian and CNS,
 CC leukemia and melanoma.
 CC
 XX Sequence 243 AA;
 SQ
 Query Match 100.0%; Score 243; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212; Indels 0; Gaps 0;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHNGSQGLPGRDGRDGDAG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHNGSQGLPGRDGRDGDAG 60
 QY 61 APGKGGGRRGLGPRGDDPPRGAGPAGTGPAGGSCVPPRASFSAKRSRSRVPSPSD 120
 DB 61 APGKGGGRRGLGPRGDDPPRGAGPAGTGPAGGSCVPPRASFSAKRSRSRVPSPSD 120
 QY 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYFFAVHATVYRASLQFDLVKNGESIASFQ 180
 DB 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYFFAVHATVYRASLQFDLVKNGESIASFQ 180
 QY 181 FFGGMPKPRASLSGGAMVRLBEPEDQWVQVGVGYTIGIYASIKTSTPSGFLVYSDMHSSP 240
 DB 181 FFGGMPKPRASLSGGAMVRLBEPEDQWVQVGVGYTIGIYASIKTSTPSGFLVYSDMHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 6
 ID AAY93688 standard; Protein; 243 AA.
 AC AAY93688;
 XX
 XX 03-OCT-2000 (first entry)
 XX
 XX Amino acid sequence of novel polypeptide PRO344.
 XX
 XX PRO201; PRO327; PRO1265; PRO344; PRO343; PRO347; PRO357;
 KW PRO715; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell;
 KM tumorigenesis; cancer; neoplastic cell growth; cell proliferation.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers

FT Peptide 1..15
 FT /note= "signal peptide"
 FT Modified-site 11..17
 FT /note= "N-myristoylation site"
 FT Modified-site 68..74
 FT /note= "N-myristoylation site"
 FT Region 77..80
 FT /note= "cell attachment sequence"
 FT Modified-site 216..222
 FT /note= "N-myristoylation site"
 EN W0200037640-A2.
 PD 29-JUN-2000.
 XX
 XX 16-DEC-1999; 99WO-US30095.
 XX
 XX 22-DEC-1998; 98US-0113296.
 PR 08-MAR-1999; 99WO-US05028.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28565.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 XX Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;
 PI Wood WI;
 XX
 DR WPI; 2000-452188/39.
 DR N-PSDB; AAA46907.
 XX
 PT New anti-polypeptide antibody useful in the treatment and diagnosis of
 PT neoplastic cell growth and proliferation
 PS Claim 61; Fig 10; 220pp; English.
 CC The present sequence represents a novel human polypeptide. The
 CC specification describes novel polypeptides designated PRO201, PRO292,
 CC PRO327, PRO1265, PRO344, PRO343, PRO347, PRO715, PRO1017,
 CC PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in
 CC the genome of tumour cells. The polypeptides are believed to contribute
 CC to tumorigenesis. The polypeptides are useful target for the
 CC identification of certain cancers, and may act as predictors of the
 CC prognosis of tumour treatment. Antibodies against these polypeptides
 CC are useful in the treatment and diagnosis of neoplastic cell growth
 CC and proliferation in mammals.
 XX
 XX Sequence 243 AA;
 SQ
 Query Match 100.0%; Score 243; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212; Indels 0; Gaps 0;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHNGSQGLPGRDGRDGDAG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHNGSQGLPGRDGRDGDAG 60
 QY 61 APGKGGGRRGLGPRGDDPPRGAGPAGTGPAGGSCVPPRASFSAKRSRSRVPSPSD 120
 DB 61 APGKGGGRRGLGPRGDDPPRGAGPAGTGPAGGSCVPPRASFSAKRSRSRVPSPSD 120
 QY 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYFFAVHATVYRASLQFDLVKNGESIASFQ 180
 DB 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYFFAVHATVYRASLQFDLVKNGESIASFQ 180
 QY 181 FFGGMPKPRASLSGGAMVRLBEPEDQWVQVGVGYTIGIYASIKTSTPSGFLVYSDMHSSP 240
 DB 181 FFGGMPKPRASLSGGAMVRLBEPEDQWVQVGVGYTIGIYASIKTSTPSGFLVYSDMHSSP 240

QY 241 VFA 243
 |||
 Db 241 VFA 243

RESULT 7
 AAB01318
 ID AAB01318 standard; Protein; 243 AA.

XX AAB01318;

DT 25-SEP-2000 (first entry)

XX Human PRO344 polypeptide.

XX PRO: membrane bound protein; secreted protein; PRO357; PRO327;
 KM PRO243; PRO715; PRO241; PRO323; PRO239; PRO344; PRO347;
 KM PRO355; PRO361; PRO365; transmembrane polypeptide;
 KM antibody; screening; detection; inhibition; probe; primer; human.
 XX Homo sapiens.

OS
 FH Key Location/Qualifiers
 FT Peptide 1..15
 FT /label= Signal peptide

FT Modified-site 11..17
 FT /note= "N-myristoylation site"

FT Modified-site 68..74
 FT /note= "N-myristoylation site"

FT Region 77..80
 FT /label= Cell attachment sequence

FT Modified-site 216..222
 FT /note= "N-myristoylation site"

FN WO20002776-A2.

XX 08-JUN-2000.

PD 01-DEC-1999; 99WO-US28301.

PF 01-DEC-1998; 98WO-US25108.

PR 16-DEC-1998; 98US-0112850.

PR 22-DEC-1998; 98US-0113296.

XX (GETH) GENENTECH INC.

PI Baker KP, Botstein D, Eaton DL, Ferrara N, Fliviaroff B;
 PI Gerltsen ME, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL;
 PI Hillan KJ, Kijavini J, Napier MA, Roy MA, Tumas D, Wood WI;

XX WPI; 2000-412324/35.

DR N-PSDB; AAA49560.

XX New human nucleic acids encoding secreted and transmembrane
 PT polypeptides, designated as PRO polypeptides, useful as pharmaceutical
 PT and diagnostic agents

PS Claim 12; Fig 18; 187pp; English.

XX New human nucleic acids encoding secreted and transmembrane
 CC polypeptides which are designated as PRO polypeptides are described
 CC The membrane-bound proteins have various industrial applications,
 CC including as pharmaceutical and diagnostic agents. The membrane-bound
 CC proteins can also be employed for screening of potential peptide or
 CC small molecule inhibitors of the relevant receptor/ligand interaction.
 CC Anti-PRO antibodies are useful for the affinity purification of PRO
 CC from recombinant cell culture or natural sources.

XX Sequence 243 AA;

Query Match 100.0%; Score 243; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4,9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHGLPCTPGHSSQGLPGDGDGDGARG 60
 |||
 Db 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHGLPCTPGHSSQGLPGDGDGDGARG 60
 QY 61 APGEKGEGRPGPLRGRGDPGRGEGAPGTGPAEGCVPRPSAFSARSSRPVPPSD 120
 |||
 Db 61 APGEKGEGRPGPLRGRGDPGRGEGAPGTGPAEGCVPRPSAFSARSSRPVPPSD 120
 QY 121 APPLPDRVLVNEQGHDAVTGKFTCOVPGVYFAVAHATYRASLQFDLVNGESIASFPQ 180
 |||
 Db 121 APPLPDRVLVNEQGHDAVTGKFTCOVPGVYFAVAHATYRASLQFDLVNGESIASFPQ 180
 QY 181 FPGGMPKPSLSCGANVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLVYSDWHSSP 240
 |||
 Db 181 FPGGMPKPSLSCGANVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLVYSDWHSSP 240
 QY 241 VFA 243
 |||
 Db 241 VFA 243

RESULT 8
 AAU12352
 ID AAU12352 standard; Protein; 243 AA.

XX AAU12352;

DT 24-OCT-2001 (first entry)

XX Human PRO344 polypeptide sequence.

XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KM breast; prostate; cervical; tumor necrosis factor-alpha; TNF-alpha;
 KM cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KM adipocyte; A-peptide; factor VIIA; gene therapy.

XX Homo sapiens.

OS WO200140466-A2.

XX 07-JUN-2001.

PD 01-DEC-2000; 2000WO-US32678.

PF 01-DEC-1999; 99WO-US28301.

PR 02-DEC-1999; 99WO-US28551.

PR 02-DEC-1999; 99WO-US28564.

PR 02-DEC-1999; 99US-0170262.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 30-DEC-1999; 99WO-US31243.

PR 06-JAN-2000; 2000WO-US00277.

PR 06-JAN-2000; 2000WO-US00376.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 24-FEB-2000; 2000WO-US04914.

PR 24-FEB-2000; 2000WO-US05004.

PR 01-MAR-2000; 2000WO-US05601.

PR 20-MAR-2000; 2000WO-US07377.

PR 21-MAR-2000; 2000WO-US07532.

PR 30-MAR-2000; 2000WO-US08439.

PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.

PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.

PR 10-NOV-2000; 2000WO-US30873.

PA (GETH) GENENTECH INC.
 XX Baker KP, Bersini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI; 2001-408281/43.
 DR N-PSDB; AAS21424.
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 XX Claim 12; Fig 362; 813pp; English.
 XX
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (BMMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 CC
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 243; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPGTRGHGSGQLPGRDGRDGDAPG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPGTRGHGSGQLPGRDGRDGDAPG 60
 QY 61 APEKEGEGRPGGLPGPRDGPGEAGPAGPTGPAGECSVPFSAFSAKRSSESRVPPSD 120
 DB 61 APEKEGEGRPGGLPGPRDGPGEAGPAGPTGPAGECSVPFSAFSAKRSSESRVPPSD 120
 QY 121 APPLPFRVLVNEQGHYDAVTKFTCOVPGVYFAVHAHTVTRASIQFLVKNGESIASFPQ 180
 DB 121 APPLPFRVLVNEQGHYDAVTKFTCOVPGVYFAVHAHTVTRASIQFLVKNGESIASFPQ 180
 QY 181 FFGGMPKPASLSGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSTSGFLVSDMHS 240
 DB 181 FFGGMPKPASLSGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSTSGFLVSDMHS 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 9
 AAB65815
 ID AAB65815 standard; Protein; 243 AA.
 XX
 AC AAB65815;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE Human TANGO 253 SEQ ID NO: 3.

XX Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
 KW INTERCEPT 258; coronary disorder; olfactory disorder;
 KW neurological disorder; pulmonary disorder; immunological disorder;
 KW developmental disorder; kidney disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200078808-A1.
 XX
 XX 28-DEC-2000.
 XX
 PF 19-JUN-2000; 2000MO-US16883.
 XX
 PR 18-JUN-1999; 99US-0336536.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 PI Leiby KR, McKay C, Bossone S;
 DR WPI; 2001-050109/06.
 XX
 XX New nucleic acids for treating diseases and disorders, e.g.
 PT atherosclerosis, infection, autoimmune diseases, obesity, ear
 PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
 PT sclerosis and asthma
 XX
 PS Claim 9; Page 211-212; 332pp; English.
 XX
 CC The present invention provides the protein and coding sequences of the
 CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
 CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
 CC coronary, pulmonary, olfactory, immunological, neurological,
 CC developmental and kidney disorders.
 CC
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 243; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPGTRGHGSGQLPGRDGRDGDAPG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPGTRGHGSGQLPGRDGRDGDAPG 60
 QY 61 APEKEGEGRPGGLPGPRDGPGEAGPAGPTGPAGECSVPFSAFSAKRSSESRVPPSD 120
 DB 61 APEKEGEGRPGGLPGPRDGPGEAGPAGPTGPAGECSVPFSAFSAKRSSESRVPPSD 120
 QY 121 APPLPFRVLVNEQGHYDAVTKFTCOVPGVYFAVHAHTVTRASIQFLVKNGESIASFPQ 180
 DB 121 APPLPFRVLVNEQGHYDAVTKFTCOVPGVYFAVHAHTVTRASIQFLVKNGESIASFPQ 180
 QY 181 FFGGMPKPASLSGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSTSGFLVSDMHS 240
 DB 181 FFGGMPKPASLSGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSTSGFLVSDMHS 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 10
 AAB49593
 ID AAB49593 standard; Protein; 243 AA.
 XX
 AC AAB49593;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Human adipocyte complement related protein homolog zs1939.
 XX
 KW Human; zacr5; gene therapy; complement inhibition; C1q domain;

KM adipocyte complement related protein homolog;
 KM inflammation; hormone secretion; inositol phosphate; arachidonate;
 KM phospholipase C activation; gastric emptying; neutrophil activation;
 KM superoxide anion production; antimicrobial; acute vascular injury;
 KM wound healing; zsig39.

OS Homo sapiens.

PN WO200073444-A1.

PD 07-DEC-2000.

PP 18-MAY-2000; 2000WO-US13608.

PR 27-MAY-1999; 99US-0321372.

PA (ZYMO) ZYMOGENETICS INC.

PI Piddington CS, Sheppard PO;

DR WPI; 2001-061531/07.

PT Novel adipocyte complement related protein homolog, ZACRP5, useful for
 PT diagnosing and treating inflammation, vascular injury microbial
 PT infections, and in wound healing

PS Disclosure; Fig 1; 121pp; English.

XX The present invention relates to human adipocyte complement related
 CC protein homolog, zacrp5 protein and coding sequence (see ABA49590 and
 CC AAC90045). ZACRP5 has a carboxyl-terminal C1q domain. The zacrp5 gene is
 CC located on human chromosome 16. zacrp5 gene and protein are useful for
 CC diagnosing and treating inflammations, for determining arterial
 CC remodeling, for modulating calcium ion concentration, hormone
 CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC zacrp5 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, for
 CC treating acute vascular injury, and for wound healing. The present
 CC sequence is human adipocyte complement related protein homolog, zsig39
 CC protein. This protein was used in a sequence homology alignment with
 CC zacrp5 protein.

XX Sequence 243 AA;

Query Match 100.0%; Score 243; DB 22; Length 243;

Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLIVLLILGLAGSPPLDNDKIPSLCPGHPGLPCTPGHHSGQLPGDGRDGDGAPG 60
 DB 1 MRPLIVLLILGLAGSPPLDNDKIPSLCPGHPGLPCTPGHHSGQLPGDGRDGDGAPG 60
 QY 61 APGEKGBGGRPGLPGRPDGPRGAGPAGTGPAGECVPRPSAFSAKRSRVPSPSD 120
 DB 61 APGEKGBGGRPGLPGRPDGPRGAGPAGTGPAGECVPRPSAFSAKRSRVPSPSD 120
 QY 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
 DB 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
 QY 181 FFGGMPKPASLSGAMVRLBPDQVWVGVDYIGIYASIKTDTSTFGFLVYSMDHSSP 240
 DB 181 FFGGMPKPASLSGAMVRLBPDQVWVGVDYIGIYASIKTDTSTFGFLVYSMDHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243

RESULT 11
 ABA49599

ID ABA49599 standard; Protein; 243 AA.

XX ABA49599;

DT 13-MAR-2001 (first entry)

DE Human adipocyte complement related protein homolog zsig39.

XX Human; zacrp6; gene therapy; complement inhibition; C1q domain;

KM adipocyte complement related protein homolog;

KM inflammation; hormone secretion; inositol phosphate; arachidonate;
 KM phospholipase C activation; gastric emptying; neutrophil activation;
 KM superoxide anion production; antimicrobial; acute vascular injury;
 KM wound healing; zsig39.

OS Homo sapiens.

PN WO200073446-A2.

PD 07-DEC-2000.

PP 22-MAY-2000; 2000WO-US14024.

PR 27-MAY-1999; 99US-0321262.

PA (ZYMO) ZYMOGENETICS INC.

PI Piddington CS, Sheppard PO;

DR WPI; 2001-061532/07.

PT Novel adipocyte complement related protein homolog, ZACRP6, useful as
 PT modulators of neurotransmission and for treating disseminated
 PT intravascular coagulation, arteriosclerosis and acute vascular injury

PS Disclosure; Fig 1; 119pp; English.

XX The present invention relates to human adipocyte complement related
 CC protein homolog, ZACRP6 protein and coding sequence (see ABA49596 and
 CC AAC90051). ZACRP6 has a carboxyl-terminal C1q domain. The zacrp6 gene is
 CC located on human chromosome 21q. zacrp6 gene and protein are useful for
 CC diagnosing and treating inflammations, for determining arterial
 CC remodeling, for modulating calcium ion concentration, hormone
 CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC zacrp6 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, for
 CC treating acute vascular injury, disseminated intravascular coagulation,
 CC arteriosclerosis and for wound healing. The present sequence is human
 CC adipocyte complement related protein homolog zsig39. This protein was
 CC used in a sequence homology comparison with ZACRP6 protein.

XX Sequence 243 AA;

Query Match 100.0%; Score 243; DB 22; Length 243;

Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLIVLLILGLAGSPPLDNDKIPSLCPGHPGLPCTPGHHSGQLPGDGRDGDGAPG 60
 DB 1 MRPLIVLLILGLAGSPPLDNDKIPSLCPGHPGLPCTPGHHSGQLPGDGRDGDGAPG 60
 QY 61 APGEKGBGGRPGLPGRPDGPRGAGPAGTGPAGECVPRPSAFSAKRSRVPSPSD 120
 DB 61 APGEKGBGGRPGLPGRPDGPRGAGPAGTGPAGECVPRPSAFSAKRSRVPSPSD 120
 QY 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
 DB 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
 QY 181 FFGGMPKPASLSGAMVRLBPDQVWVGVDYIGIYASIKTDTSTFGFLVYSMDHSSP 240
 DB 181 FFGGMPKPASLSGAMVRLBPDQVWVGVDYIGIYASIKTDTSTFGFLVYSMDHSSP 240

DB 61 EAGPAGTGPAGGECVPPKSAFSAKRSERVPPSDAPLPDRVLVNEQGHDAVTGKFT 120
QY 145 CQVGVVYFAVHAATVYRASLQFDLVKNGESIASFFQFGGMPKPSASISGAMVRLBEPDQ 204
DB 121 CQVGVVYFAVHAATVYRASLQFDLVKNGESIASFFQFGGMPKPSASISGAMVRLBEPDQ 180
QY 205 VVVGVGVDYIGIYASIKTDTSTFSGFLVYS DMH 237
DB 181 VVVGVGVDYIGIYASIKTDTSTFSGFLVYS DMH 213

RESULT 14
AAB65888
ID AAB65888 standard; Protein; 243 AA.
XX AAB65888;
AC
XX
XX 28-MAR-2001 (first entry)
DT
XX
XX Human secreted protein related protein SEQ ID NO: 102.
DE
XX
XX Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW INTERCEPT 258; coronary disorder; olfactory disorder;
KW neurological disorder; pulmonary disorder; immunological disorder;
KW developmental disorder; kidney disorder.
XX
XX Homo sapiens.
OS
XX WO200078808-A1.
PN
XX 28-DEC-2000.
PD
XX 19-JUN-2000; 2000WO-US16883.
PF
XX 18-JUN-1999; 99US-0336536.
PR
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
PI Leiby KR, McKay C, Bossone S;
XX
XX WPI; 2001-050109/06.
DR
XX
XX New nucleic acids for treating diseases and disorders, e.g.
PT atherosclerosis, infection, autoimmune diseases, obesity, ear
PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT sclerosis and asthma -
XX
XX
PS Disclosure; Page 270-271; 332pp; English.
XX
XX The present invention provides the protein and coding sequences of the
CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC coronary, pulmonary, olfactory, immunological, neurological,
CC developmental and kidney disorders.
XX
XX Sequence 243 AA;
SQ

Query Match 76.1%; Score 185; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.8e-159;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 SPVFA 243
DB 239 SPVFA 243
QY 239 SPVFA 243
DB 239 SPVFA 243

RESULT 15
AAB65891
ID AAB65891 standard; Protein; 243 AA.
XX AAB65891;
AC
XX
XX 28-MAR-2001 (first entry)
DT
XX
XX Human secreted protein related protein SEQ ID NO: 108.
DE
XX
XX Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW INTERCEPT 258; coronary disorder; olfactory disorder;
KW neurological disorder; pulmonary disorder; immunological disorder;
KW developmental disorder; kidney disorder.
XX
XX Homo sapiens.
OS
XX WO200078808-A1.
PN
XX 28-DEC-2000.
PD
XX 19-JUN-2000; 2000WO-US16883.
PF
XX 18-JUN-1999; 99US-0336536.
PR
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
PI Leiby KR, McKay C, Bossone S;
XX
XX WPI; 2001-050109/06.
DR
XX
XX New nucleic acids for treating diseases and disorders, e.g.
PT atherosclerosis, infection, autoimmune diseases, obesity, ear
PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT sclerosis and asthma -
XX
XX
PS Disclosure; Page 274; 332pp; English.
XX
XX The present invention provides the protein and coding sequences of the
CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC coronary, pulmonary, olfactory, immunological, neurological,
CC developmental and kidney disorders.
XX
XX Sequence 243 AA;
SQ

Query Match 71.6%; Score 174; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.7e-149;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHGRLGTRGHGSGQLPGRDGRDAGP 60
DB 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHGRLGTRGHGSGQLPGRDGRDAGP 60
QY 61 ARGKGSGRGLPGRPGDQPRGEGAPAGPTGAGCSPPSAFSAKRSERVPPSD 120
DB 61 ARGKGSGRGLPGRPGDQPRGEGAPAGPTGAGCSPPSAFSAKRSERVPPSD 120
QY 121 APDPFDRVLVNEQGHDAVTGKFTQVPGVYFAVHAATVYRASLQFDLVNNGES 174
DB 121 APDPFDRVLVNEQGHDAVTGKFTQVPGVYFAVHAATVYRASLQFDLVNNGES 174

Search completed: June 20, 2003, 11:40:48
Job time : 37 secs

GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 243
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-3

Query Match 100.0%; Score 243; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-218;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLTGLAAGSPPLDNDKIPSLCPGHPGLPTGHHGSGQLPGRDGRDGRDGAAG 60
DB 1 MRPLVLLTGLAAGSPPLDNDKIPSLCPGHPGLPTGHHGSGQLPGRDGRDGRDGAAG 60
QY 61 APGEGEGRRPLPGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRD 120
DB 61 APGEGEGRRPLPGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRD 120
QY 121 APLPDRVLVNEQGHYDAVTGKTCQVGVVYFAVHATVYASLQFDLVKNGESIASFFQ 180
DB 121 APLPDRVLVNEQGHYDAVTGKTCQVGVVYFAVHATVYASLQFDLVKNGESIASFFQ 180
QY 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVSDMHSSP 240
DB 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 3

US-09-336-536-4
Sequence 4, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-4

Query Match 93.8%; Score 228; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.4e-204;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SPLDNDKIPSLCPGHPGLPTGHHGSGQLPGRDGRDGRDGRDGRDGRDGRDGRDGRD 75
DB 1 SPLDNDKIPSLCPGHPGLPTGHHGSGQLPGRDGRDGRDGRDGRDGRDGRDGRDGRD 75
QY 76 PRDPPRGRBAGPAGTGTGAGGCSVPPRSGAFSAKSESRYPPSDAPLPFDRVLVNEQGH 135
DB 61 PRDPPRGRBAGPAGTGTGAGGCSVPPRSGAFSAKSESRYPPSDAPLPFDRVLVNEQGH 120

QY 136 YDAVTGKTCQVGVVYFAVHATVYASLQFDLVKNGESIASFFQFGMPKPSLSCGA 195
DB 121 YDAVTGKTCQVGVVYFAVHATVYASLQFDLVKNGESIASFFQFGMPKPSLSCGA 180
QY 196 MVRLEPDQVWVGVDYIGIYASIKTDSFGFLVSDMHSSPVA 243
DB 181 MVRLEPDQVWVGVDYIGIYASIKTDSFGFLVSDMHSSPVA 228

RESULT 4

US-09-336-536-7
Sequence 7, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 128
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-7

Query Match 52.7%; Score 128; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 9.7e-112;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 AFSAKSESRYPPSDAPLPDRVLVNEQGHYDAVTGKTCQVGVVYFAVHATVYASL 164
DB 1 AFSAKSESRYPPSDAPLPDRVLVNEQGHYDAVTGKTCQVGVVYFAVHATVYASL 60
QY 165 QFDLVKNGESIASFFQFGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTD 224
DB 61 QFDLVKNGESIASFFQFGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTD 120
QY 225 STFGFLV 232
DB 121 STFGFLV 128

RESULT 5

US-09-188-930-295
Sequence 295, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
TITLE OF INVENTION: and Methods For Their Use
FILE REFERENCE: 11000.1011C1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 295
LENGTH: 243
TYPE: PRT
ORGANISM: Rat
US-09-188-930-295

Query Match 28.8%; Score 70; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-57;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 174 SIASFQFPGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLVY 233
|||
DB 174 SIASFQFPGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLVY 233
OY 234 SDMHSSPVFA 243
|||
DB 234 SDMHSSPVFA 243

RESULT 6
US-09-336-536-11
; Sequence 11, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-11

Query Match 25.5%; Score 62; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.2e-50;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 182 FGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLVYSDMHSSPV 241
|||
DB 167 FGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLVYSDMHSSPV 226
OY 242 FA 243
||
DB 227 FA 243

RESULT 7
US-09-336-536-10
; Sequence 10, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-10

Query Match 25.5%; Score 62; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.5e-50;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 182 FGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLVYSDMHSSPV 241
|||
DB 182 FGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLVYSDMHSSPV 241
OY 242 FA 243
||
DB 242 FA 243
```

```
RESULT 8
US-09-336-536-6
; Sequence 6, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-6

Query Match 24.7%; Score 60; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.6e-49;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 GTPGHGSGILPGRDGDRDGAAPGAPGKGGKGPGLPGPRDPPGPRGAGPAGTPGA 95
|||
DB 1 GTPGHGSGILPGRDGDRDGAAPGAPGKGGKGPGLPGPRDPPGPRGAGPAGTPGA 60

RESULT 9
US-09-336-536-14
; Sequence 14, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-14

Query Match 21.0%; Score 51; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 4.4e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 182 FGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLV 232
|||
DB 78 FGGMPKASISGGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLV 128

RESULT 10
US-09-336-536-13
; Sequence 13, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
```

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 60
TYPE: PRT
ORGANISM: Mus musculus
US-09-336-536-13

Query Match 17.7%; Score 43; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.3e-33;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 36 GTPGHSGQGLPGRGDRDGPAPGAPGKGGKGGPGGPGG 78
Db 1 GTPGHSGQGLPGRGDRDGPAPGAPGKGGKGGPGGPGG 43

RESULT 11
US-09-336-536-5
Sequence 5, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT FILING DATE: 1999-06-18
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-5

Query Match 6.2%; Score 15; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 MRPLVLLLLGLAAG 15
Db 1 MRPLVLLLLGLAAG 15

RESULT 12
US-09-227-357-190
Sequence 190, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT FILING DATE: 1999-01-08
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: US/09/227,357
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 190
LENGTH: 128
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (128)
OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-190

Query Match 4.5%; Score 11; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 45 GLPGDRGRDR 55
Db 59 GLPGDRGRDR 69

RESULT 13
US-08-931-820-4
Sequence 4, Application US/08931820
Patent No. 6010863
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Assay for collagen degradation
NUMBER OF SEQUENCES: 4
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,820
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 96202596.1
FILING DATE:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1057 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Collagen type III
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1055
OTHER INFORMATION: //label= Modified
OTHER INFORMATION: /note= "Ala may be Pro"
US-08-931-820-4
Query Match 4.1%; Score 10; DB 3; Length 1057;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 GAPEKKEGG 69
Db 680 GAPEKKEGG 689
RESULT 14
US-08-963-825-21
Sequence 21, Application US/08963825
Patent No. 6110689
GENERAL INFORMATION:
APPLICANT: Qvist, Per
APPLICANT: Bonde, Martin
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and use of the Method to Diagnose the Presence of
TITLE OF INVENTION: Disorders Associated with the Metabolism of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,825

FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/187,319
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: COLLAGEN ALPHA 1 (III)
US-08-963-825-21
Query Match 4.1%; Score 10; DB 3; Length 1078;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 GAPEKKEGG 69
Db 681 GAPEKKEGG 690
RESULT 15
US-09-500-811-21
Sequence 21, Application US/09500811
Patent No. 6323314
GENERAL INFORMATION:
APPLICANT: Qvist, Per
APPLICANT: Bonde, Martin
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and use of the Method to Diagnose the Presence of
TITLE OF INVENTION: Disorders Associated with the Metabolism of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,811
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/187,319
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687

;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1078 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; IMMEDIATE SOURCE:
;; CLONE: COLLAGEN ALPHA 1 (III)
US-09-5001811-21

Query Match 4.1%; Score 10; DB 4; Length 1078;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 GAPGKGEGG 69
|||
Db 681 GAPGKGEGG 690

Search completed: June 20, 2003, 11:44:09
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:42:09 ; Search time 24 Seconds
(without alignments)
1095.594 Million cell updates/sec

Title: US-09-944-403-42
Perfect score: 243
Sequence: 1 MRPLVLLILGLAAGSPPLD.....DSTFGSLVYSDMHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 417779 seqs, 108206813 residues

Word size : 0
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppa/US07_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/prodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pep:*
10: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/prodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	100.0	243	9	US-09-944-413-42 Sequence 42, App1
2	243	100.0	243	9	US-09-944-403-42 Sequence 42, App1
3	243	100.0	243	9	US-09-944-896-42 Sequence 42, App1
4	243	100.0	243	9	US-09-944-944-42 Sequence 42, App1
5	243	100.0	243	9	US-09-944-907-42 Sequence 42, App1
6	243	100.0	243	9	US-09-944-929-42 Sequence 362, App1
7	243	100.0	243	9	US-10-028-072-362 Sequence 362, App
8	243	100.0	243	9	US-10-121-049-362 Sequence 362, App
9	243	100.0	243	9	US-10-123-904-362 Sequence 362, App
10	243	100.0	243	9	US-10-140-470-362 Sequence 362, App
11	243	100.0	243	9	US-09-796-753-68 Sequence 68, App1
12	243	100.0	243	9	US-10-176-746-362 Sequence 362, App
13	243	100.0	243	9	US-10-176-918-362 Sequence 362, App
14	243	100.0	243	9	US-10-176-921-362 Sequence 362, App
15	243	100.0	243	9	US-10-137-865-362 Sequence 362, App
16	243	100.0	243	9	US-10-140-474-362 Sequence 362, App
17	243	100.0	243	9	US-10-142-431-362 Sequence 362, App
18	243	100.0	243	9	US-10-143-114-362 Sequence 362, App
19	243	100.0	243	9	US-10-140-002-362 Sequence 362, App

20	243	100.0	243	9	US-10-142-419-362 Sequence 362, App
21	243	100.0	243	9	US-10-123-262-362 Sequence 362, App
22	243	100.0	243	9	US-10-142-423-362 Sequence 362, App
23	243	100.0	243	9	US-10-121-050-362 Sequence 362, App
24	243	100.0	243	9	US-10-141-755-362 Sequence 362, App
25	243	100.0	243	9	US-10-143-032-362 Sequence 362, App
26	243	100.0	243	9	US-10-123-108-362 Sequence 362, App
27	243	100.0	243	9	US-10-123-226-362 Sequence 362, App
28	243	100.0	243	9	US-10-123-261-362 Sequence 362, App
29	243	100.0	243	9	US-10-140-921-362 Sequence 362, App
30	243	100.0	243	9	US-10-140-928-362 Sequence 362, App
31	243	100.0	243	9	US-10-121-045-362 Sequence 362, App
32	243	100.0	243	9	US-10-123-292-362 Sequence 362, App
33	243	100.0	243	9	US-10-123-903-362 Sequence 362, App
34	243	100.0	243	9	US-10-124-819-362 Sequence 362, App
35	243	100.0	243	9	US-10-124-822-362 Sequence 362, App
36	243	100.0	243	9	US-10-140-925-362 Sequence 362, App
37	243	100.0	243	9	US-10-160-498-362 Sequence 362, App
38	243	100.0	243	9	US-09-944-884-42 Sequence 42, App1
39	243	100.0	243	9	US-10-121-041-362 Sequence 362, App
40	243	100.0	243	9	US-10-121-043-362 Sequence 362, App
41	243	100.0	243	9	US-10-121-047-362 Sequence 362, App
42	243	100.0	243	9	US-10-123-215-362 Sequence 362, App
43	243	100.0	243	9	US-10-123-902-362 Sequence 362, App
44	243	100.0	243	9	US-10-123-908-362 Sequence 362, App
45	243	100.0	243	9	US-10-123-909-362 Sequence 362, App

ALIGNMENTS

RESULT 1
US-09-944-413-42
Sequence 42, Application US/09944413
Patent No. US20020156004A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tomas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944, 413
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866, 028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067, 411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069, 334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069, 696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 694
PRIOR FILING DATE: December 16, 1997

PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020156004A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020156004A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/065520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-413-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRPLVLLILGLAAGSPPLDDNKIPSLCPGHPLPCTPGHHSQCLPRDGRDGDGARG 60
|||||

Db 1 MRPLVLLILGLAAGSPPLDDNKIPSLCPGHPLPCTPGHHSQCLPRDGRDGDGARG 60
Qy 61 ARCEKGEGRPGILPGPRGDPGRGAGPAGTGPAGCECVPPRSASFARSRSRVPSPD 120
Db 61 ARCEKGEGRPGILPGPRGDPGRGAGPAGTGPAGCECVPPRSASFARSRSRVPSPD 120
Qy 121 APLPFRVLVNEGCHDAVTGKFTCVPRGVYFAVAHTYTRASLPDLVKNESIASPFO 180
Db 121 APLPFRVLVNEGCHDAVTGKFTCVPRGVYFAVAHTYTRASLPDLVKNESIASPFO 180
Qy 181 FRGMPKPSLSGSAWVRLEPEDQWVQVGVGDYIGIYASIKTDSFSGFLYSDMHSR 240
Db 181 FRGMPKPSLSGSAWVRLEPEDQWVQVGVGDYIGIYASIKTDSFSGFLYSDMHSR 240
Qy 241 VFA 243
Db 241 VFA 243

RESULT 2
US-09-944-403-42
Sequence 42, Application US/09944403
Patent No. US20020165143A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Batou, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerltsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OR INVENTION: ACTDS ENCODING THE SAME
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,403
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998

;; PRIOR APPLICATION NUMBER: 60/074,092
;; PRIOR FILING DATE: February 9, 1998
;; PRIOR APPLICATION NUMBER: 60/075,945
;; PRIOR FILING DATE: February 25, 1998
;; PRIOR APPLICATION NUMBER: 60/112,850
;; PRIOR FILING DATE: December 16, 1998
;; PRIOR APPLICATION NUMBER: 60/113,296
;; PRIOR FILING DATE: December 22, 1998
;; PRIOR APPLICATION NUMBER: 60/146,222
;; PRIOR FILING DATE: July 28, 1999
;; PRIOR APPLICATION NUMBER: PCT/US98/19330
;; PRIOR FILING DATE: September 16, 1998
;; PRIOR APPLICATION NUMBER: PCT/US98/25108
;; PRIOR FILING DATE: December 1, 1998
;; PRIOR APPLICATION NUMBER: 09/216,021
;; PRIOR FILING DATE: December 16, 1998
;; PRIOR APPLICATION NUMBER: 09/216,517
;; PRIOR FILING DATE: December 22, 1998
;; PRIOR APPLICATION NUMBER: 09/254,311
;; PRIOR FILING DATE: March 3, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: June 22, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: September 15, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/28409
;; PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/28301
;; PRIOR FILING DATE: December 1, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: December 16, 1999
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: February 11, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: February 22, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: March 2, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: March 30, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: May 22, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/20710
;; PRIOR FILING DATE: July 28, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: December 1, 2000
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; PRIOR FILING DATE: February 28, 2001
;; NUMBER OF SEQ ID NOS: 120
;; SEQ ID NO 42
;; LENGTH: 243
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-09-944-403-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHSQGLPGRDGRGRCAG 60
DB 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHSQGLPGRDGRGRCAG 60
QY 61 APEKEKGGRPGIPGRGDPGRGEAGPAGPTGPAGSCSVPPRSASAKRSBSRVPSPD 120
DB 61 APEKEKGGRPGIPGRGDPGRGEAGPAGPTGPAGSCSVPPRSASAKRSBSRVPSPD 120
QY 121 AAPFPDRLVNEQGHDAVTKGTCCVPGVYFAVHATYRASLODLYKNGESIASFQ 180
DB 121 AAPFPDRLVNEQGHDAVTKGTCCVPGVYFAVHATYRASLODLYKNGESIASFQ 180
QY 181 FPGWKPASLSGAMVRLPEPDQWVQVGDYIGIYASIKTDSFSGFLVYSDWHSSP 240

Db 181 FPGWKPASLSGAMVRLPEPDQWVQVGDYIGIYASIKTDSFSGFLVYSDWHSSP 240
QY 241 VFA 243
Db 241 VFA 243

RESULT 3
US-09-944-896-42
Sequence 42, Application US/09944896
Patent No. US20020168715A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerltson, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kjaavin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tomas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,896
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998

PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-896-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202; Indels 0; Gaps 0;
Matches 243; Conservative 0; Mismatches 0;

QY 1 MRLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGQLPGRDGRDGDGANG 60
DB 1 MRLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGQLPGRDGRDGDGANG 60
QY 61 APGKGEGRPGIPGPRGDPGRGEGAPGPTGPGEGSCVPPRSAPSAKRSRRVPPSPD 120
DB 61 APGKGEGRPGIPGPRGDPGRGEGAPGPTGPGEGSCVPPRSAPSAKRSRRVPPSPD 120
QY 121 APPLPDRVAVNEGSHDAVTGKTCOVPGVYFAVHATYRSLQFDLVKNGESIASPFQ 180
DB 121 APPLPDRVAVNEGSHDAVTGKTCOVPGVYFAVHATYRSLQFDLVKNGESIASPFQ 180
QY 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIVASIKTDSIFSGFLVYSWMHSP 240
DB 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIVASIKTDSIFSGFLVYSWMHSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 4
US-09-944-944-42
Sequence 42, Application US/09944944
Patent No. US20020173463A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David

APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gunney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavyn, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Thmas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,944
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,686
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/216,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020173463A1ember 30, 1999

PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020173463A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-944-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAG 60
DB 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAG 60
QY 61 APGKGGGGRGRLGPRGDPGPRGAGPAGTGPAGSGSVPRSAFSAKRSRSRVPPSPD 120
DB 61 APGKGGGGRGRLGPRGDPGPRGAGPAGTGPAGSGSVPRSAFSAKRSRSRVPPSPD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYVFAVHATVYRASIQFDLVKNGESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYVFAVHATVYRASIQFDLVKNGESIASFFQ 180
QY 181 FFGGMPKPRASISGGAMVRLPEPDQVWQVGVGDYIGIYASIKTSTSGFLVYSDMHSSP 240
DB 181 FFGGMPKPRASISGGAMVRLPEPDQVWQVGVGDYIGIYASIKTSTSGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 5
US-09-944-907-42

; Sequence 42, Application US/09944907
; Publication No. US20020198147A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Batson, Dan
; APPLICANT: Ferrata, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austen
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavlin, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret

APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,907
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-907-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAG 60
DB 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAG 60
QY 61 APGKGGGGRGRLGPRGDPGPRGAGPAGTGPAGSGSVPRSAFSAKRSRSRVPPSPD 120
DB 61 APGKGGGGRGRLGPRGDPGPRGAGPAGTGPAGSGSVPRSAFSAKRSRSRVPPSPD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYVFAVHATVYRASIQFDLVKNGESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTQVPGVYVFAVHATVYRASIQFDLVKNGESIASFFQ 180
QY 181 FFGGMPKPRASISGGAMVRLPEPDQVWQVGVGDYIGIYASIKTSTSGFLVYSDMHSSP 240
DB 181 FFGGMPKPRASISGGAMVRLPEPDQVWQVGVGDYIGIYASIKTSTSGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 6
US-09-944-929-42

; Sequence 42, Application US/09944929
; Publication No. US20020197612A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Batson, Dan
; APPLICANT: Ferrata, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austen
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavlin, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,929
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT

ORGANISM: Homo Sapien
US-09-944-929-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7,9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHGSQGLRGRDGRGARG 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHGSQGLRGRDGRGARG 60
QY 61 APEKEGEGRPGLPGRGPGRGAGPAGPTGAGECSTVPSPASAKRSERVPSPD 120
DB 61 APEKEGEGRPGLPGRGPGRGAGPAGPTGAGECSTVPSPASAKRSERVPSPD 120
QY 121 APLEPRVAVNEQGHYDAVTGKFTCOVPGVYFAVAHYATYRASLQFDLVNGBSIAFPQ 180
DB 121 APLEPRVAVNEQGHYDAVTGKFTCOVPGVYFAVAHYATYRASLQFDLVNGBSIAFPQ 180
QY 181 FPGGWKPSASGSGAMVRLPEPDQVWQVGVGDYIGIYASIKTDSFSGFLVYSDMHSSP 240
DB 181 FPGGWKPSASGSGAMVRLPEPDQVWQVGVGDYIGIYASIKTDSFSGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 7
US-10-028-072-362

Sequence 362, Application US/10028072

Publication No. US20030004311A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerltsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang

TITLE OR INVENTION:

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/028, 072

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/049911

PRIOR FILING DATE: 1997-06-18

PRIOR APPLICATION NUMBER: 60/056974

PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/059113

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059115

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059117

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059122

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059184

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059352

PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 60/059836

PRIOR FILING DATE: 1997-09-24

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062285

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062287

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062814

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/062816

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063045

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063082

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/063127

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063327

PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063329

PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063550

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063561

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063704

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/063733

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/063735

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/063738

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/063755

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/064248

PRIOR FILING DATE: 1997-11-03

PRIOR APPLICATION NUMBER: 60/064809

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065846

PRIOR FILING DATE: 1997-11-17

PRIOR APPLICATION NUMBER: 60/066364

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: 60/066453

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/066511

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/069212

PRIOR FILING DATE: 1997-12-11

PRIOR APPLICATION NUMBER: 60/069278

PRIOR FILING DATE: 1997-12-11

PRIOR APPLICATION NUMBER: 60/069334

PRIOR FILING DATE: 1997-12-11

PRIOR APPLICATION NUMBER: 60/069694

PRIOR FILING DATE: 1997-12-16

PRIOR APPLICATION NUMBER: 60/072320

PRIOR FILING DATE: 1998-01-23

PRIOR APPLICATION NUMBER: 60/073612

PRIOR FILING DATE: 1998-02-04

PRIOR APPLICATION NUMBER: 60/074086

PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/074092

PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/077791

PRIOR FILING DATE: 1998-03-12

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089332
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202; Indels 0; Gaps 0;
Matches 243; Conservative 0; Mismatches 0;

QY 1 MRPLVLLGLAASPPIDNKRITSLCPGHGLPPTPCPHSGGLPGHGRDGRGARG 60
DB 1 MRPLVLLGLAASPPIDNKRITSLCPGHGLPPTPCPHSGGLPGHGRDGRGARG 60
QY 61 AFGKGEGRPGLPGRGDPGRGAGPAGTPGAGECSVPPRSARSESRRVPPSD 120
DB 61 AFGKGEGRPGLPGRGDPGRGAGPAGTPGAGECSVPPRSARSESRRVPPSD 120
QY 121 APLPFDRLVINEQGHDAVTGKFTCVPGVYTFVAHATYRRASLOFDLYNGESIASFQ 180
DB 121 APLPFDRLVINEQGHDAVTGKFTCVPGVYTFVAHATYRRASLOFDLYNGESIASFQ 180
QY 181 FFGWPKPASTSGANVRLEPEHQVWVGVGDYIGIVASIKDSTFFSGFLVYSDMHSF 240
DB 181 FFGWPKPASTSGANVRLEPEHQVWVGVGDYIGIVASIKDSTFFSGFLVYSDMHSF 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 8

US-10-121-049-362
Sequence 362, Application US/10121049
Publication No. US20030022239A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C17
CURRENT APPLICATION NUMBER: US/10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See file Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-049-362

Query Match 100.0%; Score 243; DB 9; Length 243;

Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGDGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGDGAPG 60
QY 61 APGEKGGGRPGLPGRGDPGPRGEAGPAGPTGAGCSVPPRSAPSAKRSRRVPPSD 120
DB 61 APGEKGGGRPGLPGRGDPGPRGEAGPAGPTGAGCSVPPRSAPSAKRSRRVPPSD 120
QY 121 APLPFDRLVINEQGHYDAVTGKFTCVPGVYFAVHATVYRASLOFDLVKNGESIASFPQ 180
DB 121 APLPFDRLVINEQGHYDAVTGKFTCVPGVYFAVHATVYRASLOFDLVKNGESIASFPQ 180
QY 181 FFGGMPKPASISGAMVRLPEPDQWVVOVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGGMPKPASISGAMVRLPEPDQWVVOVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 9

US-10-123-904-362
; Sequence 362, Application US/10123904
; Publication No. US20030022328A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C54
CURRENT APPLICATION NUMBER: US/10/123,904
CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-904-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGDGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGDGAPG 60
QY 61 APGEKGGGRPGLPGRGDPGPRGEAGPAGPTGAGCSVPPRSAPSAKRSRRVPPSD 120
DB 61 APGEKGGGRPGLPGRGDPGPRGEAGPAGPTGAGCSVPPRSAPSAKRSRRVPPSD 120
QY 121 APLPFDRLVINEQGHYDAVTGKFTCVPGVYFAVHATVYRASLOFDLVKNGESIASFPQ 180
DB 121 APLPFDRLVINEQGHYDAVTGKFTCVPGVYFAVHATVYRASLOFDLVKNGESIASFPQ 180

DB 121 APLPFDRLVINEQGHYDAVTGKFTCVPGVYFAVHATVYRASLOFDLVKNGESIASFPQ 180

QY 181 FFGGMPKPASISGAMVRLPEPDQWVVOVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGGMPKPASISGAMVRLPEPDQWVVOVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240

QY 241 VFA 243
DB 241 VFA 243

RESULT 10

US-10-140-470-362
; Sequence 362, Application US/10140470
; Publication No. US20030022331A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C160
CURRENT APPLICATION NUMBER: US/10/140,470
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-470-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGDGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGDGAPG 60
QY 61 APGEKGGGRPGLPGRGDPGPRGEAGPAGPTGAGCSVPPRSAPSAKRSRRVPPSD 120
DB 61 APGEKGGGRPGLPGRGDPGPRGEAGPAGPTGAGCSVPPRSAPSAKRSRRVPPSD 120
QY 121 APLPFDRLVINEQGHYDAVTGKFTCVPGVYFAVHATVYRASLOFDLVKNGESIASFPQ 180
DB 121 APLPFDRLVINEQGHYDAVTGKFTCVPGVYFAVHATVYRASLOFDLVKNGESIASFPQ 180
QY 181 FFGGMPKPASISGAMVRLPEPDQWVVOVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGGMPKPASISGAMVRLPEPDQWVVOVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 11
US-09-796-753-68

```

; Sequence 68, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 68
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-796-753-68
Query Match      100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGRDGAAG 60
|||||

```

```

Db 1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGRDGAAG 60
QY 61 APEKKEGGRRPGLPGRGDPGPRGEGAPGPTGPAECGVPPRSASAKRSERVPSPD 120
Db 61 APEKKEGGRRPGLPGRGDPGPRGEGAPGPTGPAECGVPPRSASAKRSERVPSPD 120
QY 121 APLEPFRVLYNEBGHDAVTGKFTCOVPGVYFAVATYRASLQFDLVNGESIASFFQ 180
Db 121 APLEPFRVLYNEBGHDAVTGKFTCOVPGVYFAVATYRASLQFDLVNGESIASFFQ 180
QY 181 FFGWMPKPSASLGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLVYSDWHSPP 240
Db 181 FFGWMPKPSASLGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLVYSDWHSPP 240
QY 241 VFA 243
Db 241 VFA 243
RESULT 12
US-10-175-746-362
; Sequence 362, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330301C33
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-175-746-362
Query Match      100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGRDGAAG 60
Db 1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHSGQLPGRDGRDGRDGAAG 60
QY 61 APEKKEGGRRPGLPGRGDPGPRGEGAPGPTGPAECGVPPRSASAKRSERVPSPD 120
Db 61 APEKKEGGRRPGLPGRGDPGPRGEGAPGPTGPAECGVPPRSASAKRSERVPSPD 120
QY 121 APLEPFRVLYNEBGHDAVTGKFTCOVPGVYFAVATYRASLQFDLVNGESIASFFQ 180
Db 121 APLEPFRVLYNEBGHDAVTGKFTCOVPGVYFAVATYRASLQFDLVNGESIASFFQ 180
QY 181 FFGWMPKPSASLGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLVYSDWHSPP 240
Db 181 FFGWMPKPSASLGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLVYSDWHSPP 240

```

QY 241 VFA 243
Db 241 VFA 243

RESULT 13

US-10-176-918-362
; Sequence 362, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACTS ENCODING THE SAME
; FILE REFERENCE: P330R1C382
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-918-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHHSQGLPGRDGRDGRDAG 60
Db 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHHSQGLPGRDGRDGRDAG 60
QY 61 APGEKGGRRPGLPGRGDPGRGAGPAGPTGPAECVPPRSASFARSESRRVPPSD 120
Db 61 APGEKGGRRPGLPGRGDPGRGAGPAGPTGPAECVPPRSASFARSESRRVPPSD 120
QY 121 APLEPRVLVNEQGHYDAVTGKFTCOVPGVYFAVAATYRASLOPDLVNGESIASFPQ 180
Db 121 APLEPRVLVNEQGHYDAVTGKFTCOVPGVYFAVAATYRASLOPDLVNGESIASFPQ 180
QY 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVSDWHSSP 240
Db 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVSDWHSSP 240
QY 241 VFA 243
Db 241 VFA 243

RESULT 14

US-10-176-921-362
; Sequence 362, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACTS ENCODING THE SAME
; FILE REFERENCE: P330R1C28
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-921-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHHSQGLPGRDGRDGRDAG 60
Db 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTPGHHSQGLPGRDGRDGRDAG 60
QY 61 APGEKGGRRPGLPGRGDPGRGAGPAGPTGPAECVPPRSASFARSESRRVPPSD 120
Db 61 APGEKGGRRPGLPGRGDPGRGAGPAGPTGPAECVPPRSASFARSESRRVPPSD 120
QY 121 APLEPRVLVNEQGHYDAVTGKFTCOVPGVYFAVAATYRASLOPDLVNGESIASFPQ 180
Db 121 APLEPRVLVNEQGHYDAVTGKFTCOVPGVYFAVAATYRASLOPDLVNGESIASFPQ 180
QY 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVSDWHSSP 240
Db 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVSDWHSSP 240
QY 241 VFA 243
Db 241 VFA 243

RESULT 15

US-10-137-865-362
; Sequence 362, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C154
CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-865-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHGILGTPGHGSGQLPGRDGRDGDGAPG 60
1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHGILGTPGHGSGQLPGRDGRDGDGAPG 60
DB 61 APGKRGGRGRLPGPRGDPGPRGAGPAGTGPAGGCSVPPRSAFSAKSESSESVPPSD 120
61 APGKRGGRGRLPGPRGDPGPRGAGPAGTGPAGGCSVPPRSAFSAKSESSESVPPSD 120
QY 121 APLEFDRVLVNEQGHDAVTGKFTQVPGVYFPAVHATVYRASLQFDLVKNGESIASFFQ 180
121 APLEFDRVLVNEQGHDAVTGKFTQVPGVYFPAVHATVYRASLQFDLVKNGESIASFFQ 180
DB 121 APLEFDRVLVNEQGHDAVTGKFTQVPGVYFPAVHATVYRASLQFDLVKNGESIASFFQ 180
121 APLEFDRVLVNEQGHDAVTGKFTQVPGVYFPAVHATVYRASLQFDLVKNGESIASFFQ 180
QY 181 FPGWMPKPRASISGAMVRLPEPDQWVQVGVDTIGIYASIKTDSTFSGFLVSDMHSSP 240
181 FPGWMPKPRASISGAMVRLPEPDQWVQVGVDTIGIYASIKTDSTFSGFLVSDMHSSP 240
DB 241 VFA 243
241 VFA 243

Search completed: June 20, 2003, 11:47:19
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:40:05 ; Search time 21 Seconds

(without alignments)
1112.413 Million cell updates/sec

Title: US-09-944-403-42

Perfect score: 243
Sequence: 1 MRPLVLLLLGLAAGSPPLD.....DSTFGFLVSDWHSPPVFA 243.

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0 ,

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR 73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219	90.1	219	2 T14782	hypothetical prote
2	11	4.5	289	2 T14241	hypothetical prote
3	11	4.5	369	2 S33603	surfactant protein
4	11	4.5	374	1 A42046	surfactant protein
5	11	4.5	375	1 A45225	pulmonary surfacta
6	11	4.5	1670	1 CGH038	collagen alpha 3(I
7	11	4.5	2944	2 A54849	collagen alpha 1(I
8	10	4.1	210	2 B44984	collagen - nematod
9	10	4.1	247	1 LNRBPS	pulmonary surfacta
10	10	4.1	248	1 LNRHUP	pulmonary surfacta
11	10	4.1	248	1 LNRHUP	pulmonary surfacta
12	10	4.1	248	1 LNRHUP	pulmonary surfacta
13	10	4.1	248	1 LNRHUP	pulmonary surfacta
14	10	4.1	248	1 LNRHUP	pulmonary surfacta
15	10	4.1	248	1 LNRHUP	pulmonary surfacta
16	10	4.1	248	1 LNRHUP	pulmonary surfacta
17	10	4.1	248	1 LNRHUP	pulmonary surfacta
18	10	4.1	248	1 LNRHUP	pulmonary surfacta
19	10	4.1	248	1 LNRHUP	pulmonary surfacta
20	10	4.1	248	1 LNRHUP	pulmonary surfacta
21	10	4.1	248	1 LNRHUP	pulmonary surfacta
22	10	4.1	248	1 LNRHUP	pulmonary surfacta
23	10	4.1	248	1 LNRHUP	pulmonary surfacta
24	10	4.1	248	1 LNRHUP	pulmonary surfacta
25	10	4.1	248	1 LNRHUP	pulmonary surfacta
26	10	4.1	248	1 LNRHUP	pulmonary surfacta
27	10	4.1	248	1 LNRHUP	pulmonary surfacta
28	10	4.1	248	1 LNRHUP	pulmonary surfacta
29	10	4.1	248	1 LNRHUP	pulmonary surfacta
30	10	4.1	248	1 LNRHUP	pulmonary surfacta

30	9	3.7	636	2 S41067	collagen alpha 1(I
31	9	3.7	751	2 A49874	beta-amyloid precu
32	9	3.7	765	2 S42880	amyloid precursor
33	9	3.7	964	1 CGCH2S	collagen alpha 2(I
34	9	3.7	1402	2 T46707	translational initia
35	9	3.7	1669	1 CGMS4B	collagen alpha 1(I
36	9	3.7	1752	2 A45407	collagen alpha 3(I
37	9	3.7	2288	2 T29999	hypothetical prote
38	8	3.3	46	2 T65268	collagen alpha 1(I
39	8	3.3	52	2 C83241	hypothetical prote
40	8	3.3	87	2 H31482	hypothetical prote
41	8	3.3	149	2 B82017	probable integral
42	8	3.3	149	2 B81002	conserved hypothet
43	8	3.3	164	2 S76920	hypothetical prote
44	8	3.3	170	2 S74248	antibacterial pept
45	8	3.3	170	2 I38932	antibacterial pept

ALIGNMENTS

RESULT 1

T14782 hypothetical protein DKFZp586B0621.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000

C/Accession: T14782

R/Octenwelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A/Reference number: Z18184

A/Accession: T14782

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-219 <OT>

A/Cross-references: EMBL:AL110261

A/Experimental source: adult uterus; clone DKFZp586B0621

C/Genetics:

A/Note: DKFZp586B0621.1

C/Superfamily: complement C1q carboxyl-terminal homology

Query Match 90.1% ; Score 219 ; DB 2 ; Length 219 ;
Best Local Similarity 100.0% ; Pred. No. 1.5e-199 ;
Matches 219 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

QY	25	PSLCGHRGLRCPGHHSGOGLPGDDGDGDGAGARCGEGRRGLPGRRGDPGRG	84
DB	1	PSLCGHRGLRCPGHHSGOGLPGDDGDGDGAGARCGEGRRGLPGRRGDPGRG	60
QY	85	EAGPAGPTGPAEGECVPPRSAPSAKRSRRVPPSDALPFDRLVLYNEQGHDAVYTKFT	144
DB	61	EAQPAQPTGPAEGECVPPRSAPSAKRSRRVPPSDALPFDRLVLYNEQGHDAVYTKFT	120
QY	145	CVPGVYVFAVAVATVYRASLQFDLVKNGESIASFFQFGGMPKPSASLGAMVRLBPDO	204
DB	121	CVPGVYVFAVAVATVYRASLQFDLVKNGESIASFFQFGGMPKPSASLGAMVRLBPDO	180
QY	205	VWVGVDGYIGIYASITKDSFSGFLVSDWHSPPVFA 243	
DB	181	VWVGVDGYIGIYASITKDSFSGFLVSDWHSPPVFA 219	

RESULT 2

T14241 hypothetical protein F26F12.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C/Accession: T14241

R/Wilson, R.; Bentley, D.; Gattung, S.

submitted to the EMBL Data Library, April 1996

A/Description: The sequence of C. elegans cosmid F26F12.

A/Reference number: Z21493

A/Accession: T14241

A/Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-289 <MIL>
 A:Cross-references: EMBL:U55373; PIDN:AAC25888.1; GSPDB:GN00023; CESP:F26P12.1
 A:Experimental source: strain Bristol N2; clone F26F12
 C:Genetics:
 A:Gene: CESP:F26P12.1
 A:Map position: 5
 A:Introns: 45/3
 C:Superfamily: unassigned collagens

Query Match 4.5%; Score 11; DB 2; Length 289;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PGPRGAGPAG 90
 |||||
 Db 222 PGPRGAGPAG 232

RESULT 3

S33603
 surfactant protein D - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 17-Mar-1999
 C:Accession: S33603
 R:Lim, B.L.; Lu, J.; Reid, K.B.M.
 Immunology 78, 159-165, 1993
 A>Title: Structural similarity between bovine conglutinin and bovine lung surfactant protein D
 A:Reference number: S33603; MUID:93170856; PMID:8436402
 A:Accession: S33603
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-369 <LIM>
 C:Superfamily: pulmonary surfactant protein D; C-type lectin homology
 F:248-367/Domain: C-type lectin homology <LCH>

Query Match 4.5%; Score 11; DB 2; Length 369;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDGR 55
 |||||
 Db 46 GLPGRDGRDGR 56

RESULT 4

A42046
 surfactant protein D - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A42046
 R:Shimizu, H.; Fisher, J.H.; Papst, P.; Benson, B.; Lau, K.; Mason, R.J.; Voelker, D.R.
 J. Biol. Chem. 267, 1853-1857, 1992
 A>Title: Primary structure of rat pulmonary surfactant protein D. cDNA and deduced amino acid sequence
 A:Reference number: A42046; MUID:92112913; PMID:1370483
 A:Accession: A42046
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-374 <SHI>
 A:Cross-references: GB:M01231; NID:9207035; PIDN:AAA42170.1; PID:9207036
 A:Experimental source: lung
 A>Note: sequence extracted from NCBI backbone (NCBI:76027, NCBI:76031)
 C:Superfamily: pulmonary surfactant protein D; C-type lectin homology
 F:253-372/Domain: C-type lectin homology <LCH>

Query Match 4.5%; Score 11; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDGR 55
 |||||
 Db 45 GLPGRDGRDGR 55

RESULT 5

A45225
 pulmonary surfactant protein D precursor - human
 N:Alternate names: SP-D
 C:Species: Homo sapiens (man)
 C>Date: 16-Apr-1999 #sequence_revision 16-Apr-1999 #text_change 22-Jun-1999
 C:Accession: A45225; S23434; S24555; S44420; S18382; A56776
 R:Crouch, E.; Rust, K.; Veille, R.; Donis-Keller, H.; Grosso, L.
 J. Biol. Chem. 268, 2976-2983, 1993
 A>Title: Genomic organization of human surfactant protein D (SP-D). SP-D is encoded by a single gene
 A:Reference number: A45225; MUID:93155122; PMID:8428971
 A:Accession: A45225
 A:Molecule type: DNA
 A:Residues: 1-375 <CRO>
 A:Cross-references: GB:L05483; GB:L05484; GB:L05485; NID:9292505; PIDN:AMB59450.1; PI
 A:Experimental source: placenta
 A>Note: sequence extracted from NCBI backbone (NCBI:124316)
 R:Lu, J.; Willis, A.C.; Reid, K.B.M.
 Biochem. J. 284, 795-802, 1992
 A>Title: Purification, characterization and cDNA cloning of human lung surfactant protein D
 A:Reference number: S23434; MUID:92322003; PMID:1339284
 A:Accession: S23434
 A:Molecule type: mRNA
 A:Residues: 1-30, 'T', 32-121, 'P', 123-179, 'A', 181-375 <LUU1>
 A:Cross-references: EMBL:X65018; NID:934766; PIDN:CAA46152.1; PID:934767
 A:Experimental source: lung
 A:Accession: S24555
 A:Molecule type: protein
 A:Residues: 214-234, 'X', 236, 'XX', 239-241 <LUU2>
 R:Hoppe, H.J.; Barlow, P.N.; Reid, K.B.M.
 FEBS Lett. 344, 191-195, 1994
 A>Title: A parallel three stranded alpha-helical bundle at the nucleation site of col
 A:Reference number: S44420; MUID:94244769; PMID:8187882
 A:Accession: S44420
 A:Molecule type: mRNA
 A:Residues: 202-257 <HOB>
 R:Rust, K.; Grosso, L.; Zhang, V.; Chang, D.; Persson, A.; Longmore, W.; Cai, G.Z.; C
 Arch. Biochem. Biophys. 290, 116-126, 1991
 A>Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog
 A:Reference number: S18382; MUID:91378578; PMID:1898081
 A:Accession: S18382
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 'P', 60-205, 'P', 207-374, 'HR', <RUS>
 A:Cross-references: GB:L05485; NID:9292505
 A>Note: corrections to this sequence are reported in reference A56776
 R:Crouch, E.; Persson, A.; Chang, D.
 Am. J. Pathol. 142, 241-248, 1993
 A>Title: Accumulation of surfactant protein D in human pulmonary alveolar proteinosis
 A:Reference number: A56776; MUID:93142849; PMID:8424457
 A:Accession: A56776
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 46-58, 'P', 60-62, 'E', 64-72, 223-227, 'X', 229-239, 'P', 241-245, 'X', 247-256, 'X'
 A:Cross-references: PIDN:AMB25037.1; PID:9263973; PIDN:AMB25038.1; PID:9263974
 A:Experimental source: bronchoalveolar lavage
 A>Note: sequence extracted from NCBI backbone (NCBI:123024, NCBI:123023); sequence
 C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lower
 C:Comment: This protein is synthesized by alveolar type II cells.
 C:Genetics:
 A:Gene: GDB:SFTPD, SFTPD, SP-D
 A:Cross-references: GDB:132674; OMIM:178635
 A:Map position: 10q22.2-10q23.1
 C:Superfamily: pulmonary surfactant protein D; C-type lectin homology
 C:Keywords: blocked amino end; calcium; glycoprotein; hydroxyllysine; hydroxyproline;
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-375/Product: pulmonary surfactant protein D #status predicted <MAT>
 F:21-45/Domain: non-collagenous #status predicted <NC1>
 F:46-222/Domain: collagenous #status predicted <COL>
 F:223-375/Domain: non-collagenous #status predicted <NC2>
 F:254-373/Domain: C-type lectin homology <LCH>
 F:90/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:281-373,351-365/Disulfide bonds: #status predicted

Query Match 4.5%; Score 11; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDGR 55
46 GLPGRDGRDGR 56

RESULT 6

collagen alpha 3(IV) chain precursor, long splice form - human
N:Alternate names: Goodpasture antigen, procollagen alpha 3(IV) chain long splice form
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1994 #sequence revision 03-Oct-1995 #text_change 22-Jun-1999

A:Accession: A54763; A43928; A44043; A45971; A39786
R:Matryama, M.; Leinonen, A.; Mochizuki, T.; Tly99vason, K.; Reeder, S.T.

J. Biol. Chem. 269, 23013-23017, 1994
A:Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression

A:Reference number: A54763; MUID:94364994; PMID:8083201
A:Accession: A54763

A:Molecule type: mRNA

A:Residues: 1-1670 <MAR>

A:Cross-references: GB:X80031; NID:9577563; PID:9577564
A:Experimental source: Kidney

R:Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.
J. Clin. Invest. 89, 592-601, 1992

A:Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha
A:Reference number: A43928; MUID:92147878; PMID:1737849
A:Accession: A43928

A:Molecule type: mRNA

A:Residues: 1331-1524, 'I', 1526-1670 <TUR>

A:Cross-references: GB:M81379

R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 267, 19780-19784, 1992

A:Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture
A:Reference number: A44043; MUID:93015826; PMID:1400291
A:Accession: A44043

A:Molecule type: DNA; mRNA

A:Residues: 1386-1670 <OUT>

A:Cross-references: GB:M82993; NID:9177895; PIDN:AAA21610.1; PID:9177896
A:Note: Sequence extracted from NCBI backbone (NCBI:P115597)

R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 269, 17358, 1994

A:Reference number: A44738; MUID:94274734; PMID:8006044
A:Contents: annotation; erratum; correction to intronic sequence in A44043

R:Bernal, D.; Quinones, S.; Saus, J.
J. Biol. Chem. 268, 12090-12094, 1993

A:Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.
A:Reference number: A45971; MUID:93280184; PMID:8505332
A:Accession: A45971

A:Molecule type: mRNA

A:Residues: 1427-1444 <BBR>

A:Note: Sequence extracted from NCBI backbone (NCBI:P13363); sequence incorrectly ident
R:Morrisson, K.E.; Matryama, M.; Yang-Feng, T.L.; Reeder, S.T.
Am. J. Hum. Genet. 49, 545-554, 1991

A:Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of
A:Reference number: A39786; MUID:91355570; PMID:1862840
A:Accession: A39786

A:Molecule type: mRNA

A:Residues: 1453-1593, 'A', 1595-1670 <WOR>

A:Cross-references: GB:555790; NID:9234418; PIDN:AA19637.1; PID:9234419
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.

C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope
C:Genetics:

A:Gene: GDB:COL4A3

A:Cross-references: GDB:128351; OMIM:120070

A:Map position: 2q36-2q37

A:Introns: 1385/1, 1418/1, 1488/1, 1547/2, 1585/3, 1643/2 #status incomplete

A:Note: The alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha
monomer amino-terminal domains (with disulfide and desmosome cross-links), dimer
er associations in the interrupted helical domain (with disulfide and desmosome cross

A:Function: minor structural component of extracellular basement membrane in kidn
C:Superfamily: collagen alpha 1(IV) chain

C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extr
F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <V
F:29-42/Domain: amino-terminal nonhelical, NHI <NHI>

F:43-1438/Region: interrupted helical

F:791-793/Region: cell attachment (R-G-D) motif

F:996-998/Region: cell attachment (R-G-D) motif

F:1154-1156/Region: cell attachment (R-G-D) motif

F:1306-1308/Region: cell attachment (R-G-D) motif

F:1345-1347/Region: cell attachment (R-G-D) motif

F:1432-1434/Region: cell attachment (R-G-D) motif

F:1439-1670/Domain: carboxyl-terminal nonhelical, NCI <NCI>

F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>

F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>

F:31-33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status f

F:253/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted

F:1505-1511,1616-1622/Disulfide bonds: #status predicted

F:1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

Query Match

Best Local Similarity 4.5%; Score 11; DB 1; Length 1670;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 PGLPGRDGRDGR 81
DB 1300 PGLPGRDGRDGR 1310

RESULT 7

collagen alpha 1(VII) chain precursor - human

N:Alternate names: procollagen alpha 1(VII) chain

C:Species: Homo sapiens (man)
C:Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text_change 20-Sep-1999

A:Accession: A54849; PH0844; S16316; S16328; A30296; I86686
R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.

J. Biol. Chem. 269, 20256-20262, 1994
A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1

A:Reference number: A54849; MUID:94327588; PMID:8051117
A:Accession: A54849

A:Molecule type: mRNA

A:Residues: 1-2944 <CHR>

A:Cross-references: GB:I02870; NID:9987124; PIDN:AAA75438.1; PID:9987125
R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.

Biochem. Biophys. Res. Commun. 183, 958-963, 1992
A:Title: Molecular cloning and characterization of type VII collagen cDNA.

A:Reference number: PH0844; MUID:92231902; PMID:1567409
A:Accession: PH0844

A:Molecule type: mRNA

A:Residues: 'EPR', 340-475, 'RALSTASHSTLCWARTWHPNCRSGHWTPACPCNRPASRAARAG', 524-528,
A:Cross-references: DBJ:D11152; DBJ:D13694; NID:9453698; PIDN:BA02853.1; PID:9453

A:Experimental source: keratinocyte
A:Note: The authors translated the codon ACC for residues 394 and 397 as Tyr
Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991

A:Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
A:Reference number: S16316; MUID:91334380; PMID:1871109
A:Accession: S16316

A:Molecule type: mRNA

A:Residues: 815-892, 'E', 894-1439 <PAR>
A:Cross-references: GB:M6518; GB:S49017; NID:9180914; PIDN:AAA96439.1; PID:9180915
A:Experimental source: keratinocyte

R:Gammon, W.R.; Abernethy, M.L.; Padilla, K.M.; Prisyayhan, P.S.; Cook, M.E.; Wright, J.
 J. Invest. Dermatol. 89, 691-696, 1992
 A:Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
 A:Reference number: 156328; MUID:93107742; PMID:1469284
 A:Accession: 156328
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 'EPR', 372-517, 'DV', 520-540, 'W', 542-1255 <RBS>
 A:Cross-references: GB:S51236; MID:9262308; PID:AA24637.1; PID:9262309
 R:Selzer, J.L.; Bisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.
 J. Biol. Chem. 264, 3822-3826, 1989
 A:Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagen
 A:Reference number: A30296; MUID:89139437; PMID:2537292
 A:Accession: A30296
 A:Molecule type: Protein
 A:Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 2032, 'C', 2034-2041;
 A:Note: two reported peptides cannot be reliably located
 R:Greenpan, D.S.
 Hum. Mol. Genet. 2, 273-278, 1993
 A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
 A:Reference number: 148103; MUID:93271985; PMID:8499916
 A:Accession: 148103
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 2395-2871, 'S', 2873-2944 <R2>
 A:Cross-references: GB:L06862; MID:9388713; PID:AAA89196.1; PID:9388714
 R:Christiano, A.M.; Rymann, M.; Uitto, J.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
 A:Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs
 A:Reference number: A55255; MUID:94224777; PMID:8170945
 A:Accession: A55255
 A:Comment: annotation
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 C:Genetics: ed and subsequently O-glycosylated.
 A:Gene: GDB:COL7A1; EBR1; EBD1; EB
 A:Cross-references: GDB:128750; OMIM:120120
 A:Map position: 3p21.3-3p21.3
 A:Note: defects in this gene can result in dominant and recessive dystrophic epidermolys
 A:Note: there are 118 introns
 C:Complex: type VII collagen is probably a homotrimer
 C:Function:
 A:Description: structural component of extracellular polymer associated with anchoring f
 C:Superfamily: unassigned collagens; animal knittz-type proteinase inhibitor homology; f
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>
 F:17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
 F:36-201/Domain: von Willebrand factor type A repeat homology <VMA1>
 F:231-318/Domain: fibronectin type III repeat homology <FN1>
 F:327-413/Domain: fibronectin type III repeat homology <FN2>
 F:414-502/Domain: fibronectin type III repeat homology <FN3>
 F:508-593/Domain: fibronectin type III repeat homology <FN4>
 F:598-683/Domain: fibronectin type III repeat homology <FN5>
 F:686-771/Domain: fibronectin type III repeat homology <FN6>
 F:776-864/Domain: fibronectin type III repeat homology <FN7>
 F:864-952/Domain: fibronectin type III repeat homology <FN8>
 F:954-1045/Domain: fibronectin type III repeat homology <FN9>
 F:1052-1219/Domain: von Willebrand factor type A repeat homology <VMA2>
 F:1170-1172/Region: cell attachment (R-G-D) motif
 F:1189-1253/Region: cysteine/proline-rich
 F:1254-2783/Region: interrupted helical
 F:1334-1336/Region: cell attachment (R-G-D) motif
 F:2008-2010/Region: cell attachment (R-G-D) motif
 F:2553-2555/Region: cell attachment (R-G-D) motif
 F:2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
 F:2876-2929/Domain: animal knittz-type proteinase inhibitor homology <BPI>
 F:377-786, 1109/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:2167, 2176, 2185, 2188, 2664, 2667, 2673/Modified site: 4-hydroxyproline (Pro) #status exper
 F:2625, 2631/Modified site: 5-hydroxylysine (Lys) (covalent) #status experimental
 F:2625, 2631/Binding site: carboxylate (Lys) (covalent) #status experimental
 F:2634, 2802, 2804/Disulfide bonds: interchain #status predicted

Query Match

4.5%; Score 11; DB 2; Length 2944;

Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 71 GGLPGRGDDP 81
 DB 1328 GGLPGRGDDP 1338

RESULT 8
 B44984
 collagen - nematode (Haemonchus contortus) (fragment)
 C:Species: Haemonchus contortus
 C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-May-1999
 C:Accession: B44984
 R:Shamansky, L.M.; Pratt, D.; Boisvenue, R.J.; Cox, G.N.
 Mol. Biochem. Parasitol. 37, 73-86, 1989
 A:Title: Cuticle collagen genes of Haemonchus contortus and Caenorhabditis elegans at
 A:Reference number: A44984; MUID:90136718; PMID:2615789
 A:Accession: B44984
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-210 <SHA>
 A:Cross-references: GB:U04671; GB:U04670
 C:Superfamily: unassigned collagens

Query Match 4.1%; Score 10; DB 2; Length 210;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GAPGAPGEGK 66
 DB 151 GAPGAPGEGK 160

RESULT 9
 INRBPS
 pulmonary surfactant protein A precursor - rabbit
 N:Alternate names: pulmonary surfactant 32k apoprotein; pulmonary surfactant-associat
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
 C:Accession: A29931
 R:Bogaram, V.; Qiang, K.; Mendelson, C.R.
 J. Biol. Chem. 263, 2939-2947, 1988
 A:Title: The major apoprotein of rabbit pulmonary surfactant. Elucidation of primary
 A:Reference number: A29931; MUID:88139348; PMID:2830270
 A:Accession: A29931
 A:Molecule type: mRNA
 A:Residues: 1-247 <BOG>
 A:Cross-references: GB:U03542; MID:9165705; PID:AAA31465.1; PID:9165706
 A:Note: 12-Pro was also found
 A:Note: two species of mRNA, which appear to be transcribed from a single gene, could
 A:Note: the amino end of the mature protein is blocked
 C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lower
 C:Comment: This protein is a sialoglycoprotein synthesized by alveolar type II cells.
 A:Note: The presence of calcium ions.
 C:Superfamily: mannose-binding lectin; C-type lectin homology
 C:Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; g)
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-247/Product: pulmonary surfactant protein A #status predicted <MAT>
 F:27-99/Region: collagen-like
 F:126-245/Domain: C-type lectin homology <LCH>
 F:16/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:206/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 4.1%; Score 10; DB 1; Length 247;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGDG 54
 DB 39 GLPGRDGDG 48

RESULT 10

LNHUP6

N:Alternate names: pulmonary surfactant protein A precursor (clone 6A) - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: A25720

R:Floros, J.; Steinbrink, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;

J. Biol. Chem. 261, 9029-9033, 1986

A:Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfactant

A:Reference number: A25720; MUID:86250832; PMID:3755136

A:Accession: A25720

A:Molecule type: mRNA

A:Residues: 1-248 <FLO>

A:Cross-references: GB:M13686; NID:9190669; PIND:AAA60211.1; PID:9190670

A:Note: part of the sequence was confirmed by protein sequencing

A:Note: clones corresponding to two different proteins were sequenced. Cotranslational m

C:Genetics:

A:Gene: GDB:SFTPA1; SFTPA1; SP-A; SP-A1

A:Cross-references: GDB:119593; OMIM:176630

A:Map position: 10q22-10q23

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glyco

F:1-20/Domain: signal sequence #status predicted <Sig>

F:21-248/Product: pulmonary surfactant protein A #status predicted <MAT>

F:127-246/Domain: C-type lectin homology <LCH>

F:30/33,36,42,54,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status pred

F:207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 4.1%; Score 10; DB 1; Length 248;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 45 GLPGRGRDGS 54

40 GLPGRGRDGS 49

RESULT 11

LNHUP1

N:Alternate names: pulmonary surfactant protein A precursor (clone 1A) - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: B25720

R:Floros, J.; Steinbrink, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;

J. Biol. Chem. 261, 9029-9033, 1986

A:Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfactant

A:Reference number: A25720; MUID:86250832; PMID:3755136

A:Accession: B25720

A:Molecule type: mRNA

A:Residues: 1-248 <FLO>

A:Cross-references: GB:K03475

A:Note: part of the sequence was confirmed by protein sequencing

A:Note: clones corresponding to two different proteins were sequenced. Cotranslational m

C:Genetics:

A:Gene: GDB:SFTPA1; SFTPA1; SP-A; SP-A1

A:Cross-references: GDB:119593; OMIM:176630

A:Map position: 10q22-10q23

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glyco

F:1-20/Domain: signal sequence #status predicted <Sig>

F:21-248/Product: pulmonary surfactant protein A #status predicted <MAT>

F:127-246/Domain: C-type lectin homology <LCH>

F:30/33,36,42,54,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status pred

F:207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 4.1%; Score 10; DB 1; Length 248;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 45 GLPGRGRDGS 54

40 GLPGRGRDGS 49

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRGRDGS 54

Db 40 GLPGRGRDGS 49

RESULT 12

LNDRPS

N:Alternate names: pulmonary surfactant protein A precursor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999

C:Accession: A25296; A61227; A60142

R:Benson, B.; Hawgood, S.; Schilling, J.; Clements, J.; Damm, D.; Cordell, B.; Whit

Proc. Natl. Acad. Sci. U.S.A. 82, 6379-6383, 1985

A:Title: Structure of canine pulmonary surfactant apoprotein: cDNA and complete ami

A:Reference number: A25296; MUID:86016705; PMID:3863100

A:Accession: A25296

A:Molecule type: mRNA

A:Residues: 1-248 <BEN>

A:Note: the authors translated the codon TTG for residue 60 as Pro

A:Note: part of the sequence, including the amino end of the mature protein, was co

R:Liou, D.F.; Ryan, S.F.

Chem. Phys. Lipids 59, 29-38, 1991

A:Title: Purification of surfactant protein A from dog lung by reconstitution with

A:Reference number: A61227; MUID:92163993; PMID:1790579

A:Accession: A61227

A:Molecule type: protein

A:Residues: 18-32 <LTA>

R:Roos, G.P.; Meuth, J.; Ohning, B.; Kim, Y.; Whitsett, J.A.

Biochim. Biophys. Acta 870, 267-278, 1986

A:Title: Purification of canine surfactant-associated glycoproteins A. Identificati

A:Reference number: A60142; MUID:86159848; PMID:3006781

A:Accession: A60142

A:Molecule type: protein

A:Residues: 24-34;95-101, 'X', 103-108 <ROS>

R:Parthy, L.

Nature 325, 490, 1987

A:Reference number: A93388; MUID:87115834; PMID:3808053

A:Contents: annotation; animal lectin domain homology

C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that low

pendent on the presence of calcium ions.

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: alveolar proteinosis; calcium; disulfide bond; gaseous exchange; hydrox

F:1-17/Domain: signal sequence #status predicted <Sig>

F:18-248/Product: pulmonary surfactant protein A #status experimental <MPT>

F:28-102/Region: collagen-like

F:127-246/Domain: C-type lectin homology <LCH>

F:20/207/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:30/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match

Best Local Similarity 4.1%; Score 10; DB 1; Length 248;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 45 GLPGRGRDGS 54

40 GLPGRGRDGS 49

RESULT 13

LNDRPS

N:Alternate names: pulmonary surfactant protein A precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: A29299; J050034; S23183

R:Sano, K.; Fisher, J.; Mason, R.J.; Kuroki, Y.; Schilling, J.; Benson, B.; Voelker

Biochem. Biophys. Res. Commun. 144, 367-374, 1987

A:Title: Isolation and sequence of a cDNA clone for the rat pulmonary surfactant-as

A:Reference number: A29299; MUID:87213191; PMID:3579914

A:Contents: 0.9 kb cDNA
 A:Accession: A29299
 A:Molecule type: mRNA
 A:Residues: 1-248 <SN>
 A>Note: part of the sequence, including the amino end of the mature protein, was confirmed by Fisher, J.H.; Emrie, P.A.; Shannon, J.; Sano, K.; Hattler, B.; Mason, R.J. Biochem. Biophys. Acta 950, 338-345, 1988
 A>Title: Rat pulmonary surfactant protein A is expressed as two differently sized mRNA
 A:Reference number: J50034; MUID:89000785; PMID:2901856
 A:Contents: 1.6 kb cDNA
 A:Accession: J50034
 A:Molecule type: mRNA
 A:Residues: 1-248 <F16>
 A>Note: the codons given for residues 78, 84, and 180 are inconsistent with the authors' Ribicaz-Masmonelli, T.; Fraston, C.; Bourdon, J.; Raymondjean, M.; Kahn, A. Eur. J. Biochem. 206, 613-623, 1992
 A>Title: Characterization of the rat pulmonary surfactant protein A promoter.
 A:Reference number: S23183; MUID:92298987; PMID:1606951
 A:Accession: S23183
 A:Molecule type: DNA
 A:Residues: 1-32 <IAC>
 C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers the surface tension of the lung. This protein is a sialoglycoprotein synthesized by alveolar type II cells. It is dependent on the presence of calcium ions.
 C:Comment: Two species of mRNA, which probably are transcribed from a single gene, have been identified. Size heterogeneity of these proteins arises from posttranslational modification.
 C:Superfamily: mannose-binding lectin; C-type lectin homology
 C:Keywords: alveolar proteinosis; calcium; gaseous exchange; glycoprotein; hydroxyproline
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:37-109/Product: pulmonary surfactant protein A #status experimental, <MAT>
 F:127-246/Region: collagen-like
 F:21/Binding site: carbonyl site: (Asn) (covalent) #status absent
 F:30,33,36,42,54,57,63,70,76/Modified site: 4-hydroxyproline (Pro) #status: experiment
 F:207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.1%; Score 10; DB 1; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDG 54
 |||||
 DB 40 GLPGRDGRDG 49

RESULT 14
 151921
 pulmonary surfactant-associated protein A1 - human
 N:Alternate names: SP-A1
 C:Species: Homo sapiens (man)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
 C:Accession: 151921
 R:Katyai, S.L.; Singh, G.; Locker, J. Am. J. Respir. Cell Mol. Biol. 6, 446-452, 1992
 A>Title: Characterization of a second human pulmonary surfactant-associated protein SP-A
 A:Reference number: 151921; MUID:92198680; PMID:1372511
 A:Accession: 151921
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-248 <RES>
 A:Cross-references: GB:M68519; NID:G338048; PIDN:AAA60319.1; PID:G338049
 C:Genetics:
 A:Gene: GDB:SFTPA1; SFTPA1; SP-A; SP-A1
 A:Cross-references: GDB:119593; OMIM:178630
 A:Map position: 10q22-10q23
 A:Introns: 58/1; 98/1; 124/1
 C:Superfamily: mannose-binding lectin; C-type lectin homology
 F:127-246/Domain: C-type lectin homology <LCH>

Query Match 4.1%; Score 10; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDG 54
 |||||
 DB 40 GLPGRDGRDG 49

RESULT 15
 A48853
 pulmonary surfactant-associated protein SP-A - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
 C:Accession: A48853
 R:Korfhagen, T.R.; Bruno, M.D.; Glasser, S.W.; Ciruolo, P.J.; Whitsett, J.A.; Lattie
 Am. J. Physiol. 263, L546-L554, 1992
 A>Title: Murine pulmonary surfactant SP-A gene: cloning, sequence, and transcription
 A:Reference number: A48853; MUID:93072386; PMID:1443158
 A:Accession: A48853
 A:Molecule type: preliminary
 A:Residues: 1-248 <KOR>
 A:Cross-references: GB:S48768; NID:G260452; PIDN:AMB24274.1; PID:G260453
 A>Note: sequence extracted from NCBI backbone (NCBIN:118740, NCBI:P:118741)
 C:Superfamily: mannose-binding lectin; C-type lectin homology
 F:127-246/Domain: C-type lectin homology <LCH>

Query Match 4.1%; Score 10; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDG 54
 |||||
 DB 40 GLPGRDGRDG 49

Search completed: June 20, 2003, 11:43:37
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:26:54 ; Search time 17 Seconds
(without alignments)
592.868 Million cell updates/sec

Title: US-09-944-403-42

Sequence: 1 MRPLVLLGLGLAAGSPPLD.....DSTFGSLVYSDMHSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	243	100.0	243 1	Q0bXJ0 homo sapien
2	11	4.5	289 1	Q0bXJ2 homo sapien
3	11	4.5	369 1	P35246 bos taurus
4	11	4.5	374 1	P50404 mus musculus
5	11	4.5	374 1	P35248 rattus norv
6	11	4.5	375 1	P35247 homo sapien
7	11	4.5	1355 1	CA21_HUMAN
8	11	4.5	1670 1	CA34_HUMAN
9	11	4.5	2944 1	CA17_HUMAN
10	10	4.1	210 1	CAC2_HARECO
11	10	4.1	247 1	PSPA_CAVPO
12	10	4.1	247 1	PSPA_RABIT
13	10	4.1	248 1	PSPA_CANPA
14	10	4.1	248 1	PSPA_HUMAN
15	10	4.1	248 1	PSPA_MOUSE
16	10	4.1	248 1	PSPA_RAT
17	10	4.1	249 1	PSPA_PIG
18	10	4.1	525 1	CA21_RABIT
19	10	4.1	547 1	CAFI_EPHMU
20	10	4.1	1049 1	CA13_BOVIN
21	10	4.1	1464 1	CA13_MOUSE
22	10	4.1	1466 1	CA13_HUMAN
23	9	3.7	265 1	CTRC_NEIMA
24	9	3.7	636 1	CA13_RAT
25	9	3.7	695 1	APP2_MOUSE
26	9	3.7	765 1	APP2_RAT
27	9	3.7	1364 1	CA21_BOVIN
28	9	3.7	1366 1	CA21_CANPA
29	9	3.7	1402 1	IF4G_RABIT
30	9	3.7	1669 1	CA14_MOUSE
31	8	3.3	28 1	CIQC_MOUSE
32	8	3.3	170 1	FA39_HUMAN
33	8	3.3	193 1	CERB_HUMAN

34	8	3.3	193 1	CERB_MOUSE	Q9r171 mus musculus
35	8	3.3	201 1	CERL_HUMAN	Q9r177 homo sapien
36	8	3.3	224 1	CERL_RAT	P98087 rattus norv
37	8	3.3	244 1	APM1_HUMAN	Q15848 homo sapien
38	8	3.3	247 1	APM1_MOUSE	Q60994 mus musculus
39	8	3.3	301 1	CC02_CABEL	P17656 caenorhabdi
40	8	3.3	306 1	CC40_CABEL	P14804 caenorhabdi
41	8	3.3	316 1	CC12_CABEL	P20630 caenorhabdi
42	8	3.3	316 1	CC13_CABEL	P20631 caenorhabdi
43	8	3.3	329 1	CQT4_HUMAN	Q9bXJ3 homo sapien
44	8	3.3	356 1	VP36_CANPA	P49256 canis fami1
45	8	3.3	358 1	VP36_MOUSE	Q9bXh5 mus musculus

ALIGNMENTS

RESULT 1
ID COTS_HUMAN STANDARD; PRT; 243 AA.
AC Q0bXJ0; Q9UFY4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement-clg tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P.O., Humes J.M.;
RT "Homo sapiens complement-clg tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF329841; AKK17965.1; -;
DR EMBL; AL110261; CAB53702.1; -;
DR Genew; HGNC:14344; C1QTNF5.
DR InterPro; IPR001073; C1Q
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1Q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; signal.
FT SIGNAL; 1 15
FT CHAIN; 16 243
FT FT
FT DOMAIN; 30 95
FT DOMAIN; 97 243
FT FT
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDAB7EB784 CRC64;
Query Match 100.0%; Score 243; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.9e-193;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAASPPDLDNKKI PSLCPGHGPGLPCTPGHNGSGQLPGRDGRDGRDAPG 60
 DB 1 MRPLVLLILGLAASPPDLDNKKI PSLCPGHGPGLPCTPGHNGSGQLPGRDGRDGRDAPG 60
 QY 61 APEKGGKGRPGLPGRPGDGRPGEGAPPTGAPGCSVPSPSAPSAKSESRRVPPSPD 120
 DB 61 APEKGGKGRPGLPGRPGDGRPGEGAPPTGAPGCSVPSPSAPSAKSESRRVPPSPD 120
 QY 121 APLPFDRLVNEQGHDAVTKFTCCQVPGVYFAVHATVYRASLOEDLVNKGESIASFPQ 180
 DB 121 APLPFDRLVNEQGHDAVTKFTCCQVPGVYFAVHATVYRASLOEDLVNKGESIASFPQ 180
 QY 181 FPGGMPFAPSLSGAMVRLPEPDQWVQVGVGYIGIYASIKTDSFTSGFLVYSMDHSSP 240
 DB 181 FPGGMPFAPSLSGAMVRLPEPDQWVQVGVGYIGIYASIKTDSFTSGFLVYSMDHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243

RESULT 2
 COT7_HUMAN
 ID COT7_HUMAN STANDARD; PRT; 289 AA.
 AC 09BKX2;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Complement-c1g tumor necrosis factor-related protein 7 precursor.
 GN C10MP7 OR CTRP7.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RA Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.;
 RT "Homo sapiens complement-c1g tumor necrosis factor-related protein.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN 12)
 RP SEQUENCE FROM N.A.
 RC Straussberg R.;
 RA TISSUE-TESTES;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
 CC 1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: AF229839; AAK17963.1; -
 CC EMBL: BC022187; AAH22187.1; -
 CC EMBL: HGNC:14342; C1QTNF7.
 CC InterPro: IPR001073; C1Q.
 CC InterPro: IPR000087; Collagen.
 CC Pfam: PF00386; C1q; 1.
 CC Pfam: PF01391; Collagen; 2.
 CC PRINTS: PR00007; COMPLEMENTC1Q.
 CC SMART: SM00110; C1Q; 1.
 CC PROSITE: PS01113; C1Q; 1.
 CC Collagen; Signal.
 KW CHAIN 1 16
 FT SIGNAL 1 16
 FT CHAIN 17 289
 FT DOMAIN 38 139
 FT DOMAIN 141 276
 FT SEQUENCE 289 AA; 30683 MW; A61609P86D26946 CRC64;

Query Match 4.5%; Score 11; DB 1; Length 289;
 Best Local Similarity 100.0%; Pred. No. 0.06;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGR 55
 DB 59 GLPGRDGR 69

RESULT 3
 PSPD_BOVIN
 ID PSPD_BOVIN STANDARD; PRT; 369 AA.
 AC P3246;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
 GN SFTPD OR SFTPD4.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 208-247.
 RA MEDLINE=93170856; PubMed=8436402;
 RP TISSUE=Lung;
 RX Lim B.L., Lu J., Reid K.B.M.;
 RT "Structural similarity between bovine conglutinin and bovine lung
 RT surfactant protein D and demonstration of liver as a site of
 RT synthesis of conglutinin.";
 RL Immunology 78:159-165(1993).
 CC 1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
 CC MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
 CC EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE
 CC EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
 CC 1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
 CC 1- SUBCELLULAR LOCATION: Extracellular.
 CC MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
 CC PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
 CC CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
 CC HYDROPHOBIC PROTEINS (SP-B AND SP-C).
 CC 1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
 CC 1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: X75911; CAA53510.1; -
 CC PIR: S33603; S33603.
 CC HSP: P3247; 1B08.
 CC InterPro: IPR000087; Collagen.
 CC InterPro: IPR001304; lectin_C.
 CC Pfam: PF00059; lectin_C; 1.
 CC Pfam: PF01391; Collagen; 2.
 CC SMART: SM00034; CLECT; 1.
 CC PROSITE: PS00615; C-TYPE LECTIN 1; 1.
 CC PROSITE: PS00641; C-TYPE LECTIN 2; 1.
 CC Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
 KW Signal; Lectin; Collagen; Repeat; Coiled coil.
 KW Signal; 1 20
 FT SIGNAL 1 20
 FT CHAIN 21 369
 FT FT 21 369
 FT FT 21 369
 FT DOMAIN 46 216
 FT DOMAIN 217 248
 FT DOMAIN 273 369
 FT DISULFID 275 367
 D COLLAGEN-LIKE.
 D COILED COIL (POTENTIAL).
 D C-TYPE LECTIN (SHORT FORM).
 BY SIMILARITY.

FT DISULFID 345 359 BY SIMILARITY.
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 78 78 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 87 87 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 96 96 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 99 99 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 165 165 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 171 171 HYDROXYLATION (BY SIMILARITY).
 SQ SEQUENCE 369 AA; 37361 MW; 07D8B824E0AEB23 CRC64;
 Query Match 4.5%; Score 11; DB 1; Length 369;
 Best Local Similarity 100.0%; Pred. No. 0.073;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 45 GLPGRDGRDGR 55
 Db 46 GLPGRDGRDGR 56
 RESULT 4
 PSPD MOUSE STANDARD; PRT; 374 AA.
 AC P50404;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
 GN SFTPD OR SFTPA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
 RX MEDLINE=96094460; PubMed=7499852;
 RA Motwani M., White R.A., Guo N., Dowler L.L., Tauber A.I., Sattray K.N.;
 RT "Mouse surfactant protein-D. cDNA cloning, characterization, and gene
 RT localization to chromosome 14.";
 RT J. Immunol. 155:5671-5677(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=99244602; PubMed=10226065;
 RA Lawson P.R., Perkins V.C., Holmskov U., Reid K.B.;
 RT "Genomic organization of the mouse gene for lung surfactant protein
 RT D.";
 RT Am. J. Respir. Cell Mol. Biol. 20:953-963(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Fisher J.H., Shetfelyevich V.V.;
 RT "Surfactant protein-D regulates surfactant phospholipid homeostasis in
 RT vivo.";
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
 CC MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
 CC EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE
 CC EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
 CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTIMERS.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
 CC PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
 CC CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-B AND SP-C).
 CC HYDROPHOBIC PROTEINS (SP-A AND SP-D) AND 2 SMALL
 CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; LA0156; AAA92021.1; -
 DR EMBL; AF047742; AAD31380.1; -
 DR EMBL; AF047741; AAD31380.1; JOINED.
 DR EMBL; AF192134; AAF15277.1; -
 DR EMBL; BC003705; AAH03705.1; -
 DR HSSP; P35247; 1808
 DR MGD; MG1:109515; Sftpd.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001304; Lectin C.
 DR Pfam; PF00059; Lectin C; 1.
 DR Pfam; PF01391; Collagen; 3.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
 DR PROSITE; PS00641; C-TYPE LECTIN_2; 1.
 KW Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
 KW Signal; Lectin; Collagen; Repeat; Coiled coil.
 FT SIGNAL 1 19
 FT CHAIN 20 374
 FT DOMAIN 45 221
 FT DOMAIN 222 253
 FT DOMAIN 278 374
 FT DISULFID 280 372
 FT DISULFID 350 364
 FT CARBOHYD 89 89
 SQ SEQUENCE 374 AA; 37688 MW; FE034261265F43B4 CRC64;
 Query Match 4.5%; Score 11; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 45 GLPGRDGRDGR 55
 Db 45 GLPGRDGRDGR 55
 RESULT 5
 PSPD RAT STANDARD; PRT; 374 AA.
 AC P35246;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D)
 DE (CP4).
 GN SFTPD OR SFTPA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-33.
 RC TISSUE=Lung;
 RX MEDLINE=92112913; PubMed=1370483;
 RA Shimizu H., Fisher J.H., Papst P., Benson B., Lau K., Mason R.J.,
 RA Voelker D.R.;
 RT "Primary structure of rat pulmonary surfactant protein D. cDNA and
 RT deduced amino acid sequence.";
 RT J. Biol. Chem. 267:1853-1857(1992).
 RN [2]
 RP SEQUENCE OF 73-95 AND 153-180.
 RC TISSUE=Lung;
 RX MEDLINE=90001186; PubMed=2675969;
 RA Persson A., Chang D., Rust K., Moxley M., Longmore W., Crouch B.;
 RT "Purification and biochemical characterization of CP4 (SP-D), a
 RT collagenous surfactant-associated protein.";
 RT Biochemistry 28:6361-6367(1989).

FT SIGNAL 1 20 BY SIMILARITY.
 FT CHAIN 21 375 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
 FT DOMAIN 46 222 D.
 FT DOMAIN 223 252 COLLAGEN-LIKE
 FT DISULFID 279 375 COILED COIL (POTENTIAL).
 FT DISULFID 351 365 C-TYPE LECTIN (SHORT FORM).
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 78 78 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 87 87 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 96 96 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 99 96 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 171 171 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 177 177 HYDROXYLATION (BY SIMILARITY).
 FT CONFLICT 31 31 M -> T (IN REF. 2).
 FT CONFLICT 59 59 P -> F (IN REF. 3).
 FT CONFLICT 122 122 A -> P (IN REF. 2).
 FT CONFLICT 180 180 T -> A (IN REF. 2).
 FT CONFLICT 206 206 D -> P (IN REF. 3).
 FT CONFLICT 374 374 E -> EH (IN REF. 3).
 SQ SEQUENCE 375 AA; 37702 MW; 2986B2699FC01A6A CRC64;
 Query Match 4.5%; Score 11; DB 1; Length 375;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 45 GLPGRDGRDGR 55
 DB 46 GLPGRDGRDGR 56
 RESULT 7
 ID CA21_RANCA STANDARD; PRT; 1355 AA.
 AC 042350;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Collagen alpha 2(I) chain precursor.
 GN COL1A2.
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Tail;
 RX MEDLINE=97417499; PubMed=9272872;
 RA Asahina K., Oofusa K., Obata M., Yoshizato K.;
 RT "Cloning and characterization of the full length cDNA encoding alpha2
 RT type I collagen of bullfrog Rana catesbeiana."
 RL Gene 194:283-289(1997).
 CC -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
 CC (FIBRILLAR FORMING COLLAGEN).
 CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
 CC -1- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
 CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
 CC HYDROXYAPATITE.
 CC -1- PM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, D88764; BAA22380.1;
 DR InterPro; IPR00087; Collagen.
 DR InterPro; IPR000885; Fib_collagen_C.

DR Pfam; PF01391; Collagen; 18.
 DR Pfam; PF01410; COLFI; 1.
 DR Prodom; PD000007; Collagen; 3.
 DR Prodom; PD002078; Fib_collagen_C; 1.
 DR SMART; SM00038; COLFI.1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 82 AMINO-TERMINAL PROPEPTIDE (POTENTIAL).
 FT CHAIN 83 1093 COLLAGEN ALPHA 2(I) CHAIN.
 FT PROPEP 1094 1355 CARBOXYL-TERMINAL PROPEPTIDE
 FT (BY SIMILARITY).
 FT CARBOHYD 1206 1206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1256 1256 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1355 AA; 127643 MW; C8793AD5D6F41D2A CRC64;
 Query Match 4.5%; Score 11; DB 1; Length 1355;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 83 RGEAGPAGPTG 93
 DB 705 RGEAGPAGPTG 715
 RESULT 8
 ID CA34_HUMAN STANDARD; PRT; 1670 AA.
 AC 001955; Q9B0T2;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Collagen alpha 3(IV) chain precursor (Goodpasture antigen).
 GN COL4A3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=94364994; PubMed=8083201;
 RA Mariyama M., Leinonen A., Mochizuki T., Trygvasen K., Reeders S.T.;
 RT "Complete primary structure of the human alpha 3(IV) collagen chain.
 RT Coexpression of the alpha 3(IV) and alpha 4(IV) collagen chains in
 RT human tissues."
 RL J. Biol. Chem. 269:23013-23017(1994).
 RN [2]
 RP REVISIONS.
 RA Leinonen A.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A., VARIANTS AS E-297; R-407; R-640; R-1167; E-1207;
 RP O-1215; S-1277; T-1330; E-1334; E-1347 AND C-1661, AND VARIANTS R-43;
 RP E-162; Y-326; H-408; R-451; L-574; E-1269 AND P-1474.
 RX MEDLINE=21064696; PubMed=11134255;
 RA Heider L., Attwood C., Forestier L., Cohen-Solal L., Mollet G.,
 RA Gutierrez B., Stavrou C., Gubler M.C., Antignac C.;
 RT "Structure of the human type IV collagen gene COL4A3 and mutations in
 RT autosomal Alport syndrome."
 RL J. Am. Soc. Nephrol. 12:97-106(2001).
 RN [4]
 RP SEQUENCE OF 1386-1670 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93015826; PubMed=1400291;
 RA Quinones S., Bernal D., Garcia-Sogo M., Elena S.F., Saus J.;
 RT "Exon/Intron structure of the human alpha 3(IV) gene encompassing the
 RT Goodpasture antigen (alpha 3(IV)NC1). Identification of a potentially
 RT antigenic region at the triple helix/NC1 domain junction."
 RL J. Biol. Chem. 267:19780-19784(1992).
 RN [5]
 RP SEQUENCE OF 1453-1670 FROM N.A.
 RX MEDLINE=91353570; PubMed=1882840;
 RA Morrison K.E., Mariyama M., Yang-Feng T.L., Reeders S.T.;

RT "Sequence and localization of a partial cDNA encoding the human alpha
RT 3 chain of type IV collagen.";
RL Am. J. Hum. Genet. 49:545-554(1991).
RN [6]
RP SEQUENCE OF 1331-1670 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92147878; PubMed=1737849;
RA Turner N., Mason F.J., Brown R., Fox M., Povey S., Rees A.,
RA Pusey C.D.;
RT "Molecular cloning of the human Goodpasture antigen demonstrates it
RT to be the alpha 3 chain of type IV collagen.";
RL J. Clin. Invest. 89:592-601(1992).
RN [7]
RP SEQUENCE OF 1644-1670 FROM N.A.
RC TISSUE=Kidney;
RA Ding J.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 1439-1670, AND ALTERNATIVE SPLICING.
RC TISSUE=Kidney;
RX MEDLINE=94124597; PubMed=8294492;
RA Feng L., Xia Y., Wilson C.B.;
RT "Alternative splicing of the NCI domain of the human alpha 3(IV)
RT collagen gene. Differential expression of mRNA transcripts that
RT predict three protein variants with distinct carboxyl regions.";
RL J. Biol. Chem. 269:2342-2348(1994).
RN [9]
RP SEQUENCE OF 1-29, FROM N.A.
RX MEDLINE=98196854; PubMed=9537506;
RA Momota R., Sugimoto M., Ohashi T., Kigasawa K., Yoshioka H.,
RA Nimomiya Y.;
RT "Two genes, COL4A3 and COL4A4 coding for the human alpha3(IV) and
RT alpha4(IV) collagen chains are arranged head-to-head on chromosome
RT 2q36.";
RL FEBS Lett. 424:11-16(1998).
RN [10]
RP ALTERNATIVE SPLICING.
RX MEDLINE=93280184; PubMed=8505332;
RA Bernal D., Quiñones S., Saus J.;
RT "The human mRNA encoding the Goodpasture antigen is alternatively
RT spliced.";
RL J. Biol. Chem. 268:12090-12094(1993).
RN [11]
RP VARIANT PRO-1474.
RX MEDLINE=95078827; PubMed=7987301;
RA Lemmink H.H., Mochizuki T., van den Heuvel L.P.W.J., Schroeder C.H.,
RA Barentzen A., Monnens L.A.H., van Oost B.A., Brunner H.G.,
RT "Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal
RT recessive Alport syndrome.";
RL Hum. Mol. Genet. 3:1269-1273(1994).
CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
CC MESANGIAL BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
CC MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCAN AND ENACTIN/
CC NIDOCEN.
CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV) -
CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC -1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS, 1 (SHOWN HERE), 2/V AND
CC 3/L5; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER IN THEIR
CC C-TERMINAL NCI DOMAINS.
CC -1- TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE
CC COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE,
CC COCHLEA, LUNG AND BRAIN.
CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC DOMAIN (NC1) AT THEIR C-TERMINUS. FREQUENT INTERRUPTIONS OF THE
CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC TRIPLE-HELICAL 7S DOMAIN.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPLE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- PTM: THE ALTERNATIVE SPLICED FORM V CONTAINS AN ADDITIONAL

CC N-LINKED GLYCOSYLATION SITE.
CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
CC THESE, LOCATED IN THE NCI DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
CC IV COLLAGENS.
CC -1- PTM: Phosphorylated by the Goodpasture antigen-binding protein.
CC DISEASE: ANTIBODIES AGAINST THE NCI DOMAIN OF ALPHA3(IV) MEDIAE
CC THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS
CC CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.
CC -1- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I
CC AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY
CC GLOMERULONEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE,
CC HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN
CC MALES AND FEMALES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X80031; CAA56335.1; -
DR EMBL; AJ288487; CAC36101.1; JOINED.
DR EMBL; AJ288488; CAC36101.1; JOINED.
DR EMBL; AJ288489; CAC36101.1; JOINED.
DR EMBL; AJ288490; CAC36101.1; JOINED.
DR EMBL; AJ288491; CAC36101.1; JOINED.
DR EMBL; AJ288492; CAC36101.1; JOINED.
DR EMBL; AJ288493; CAC36101.1; JOINED.
DR EMBL; AJ288494; CAC36101.1; JOINED.
DR EMBL; AJ288495; CAC36101.1; JOINED.
DR EMBL; AJ288496; CAC36101.1; JOINED.
DR EMBL; AJ288497; CAC36101.1; JOINED.
DR EMBL; AJ288498; CAC36101.1; JOINED.
DR EMBL; AJ288499; CAC36101.1; JOINED.
DR EMBL; AJ288500; CAC36101.1; JOINED.
DR EMBL; AJ288501; CAC36101.1; JOINED.
DR EMBL; AJ288502; CAC36101.1; JOINED.
DR EMBL; AJ288503; CAC36101.1; JOINED.
DR EMBL; AJ288504; CAC36101.1; JOINED.
DR EMBL; AJ288505; CAC36101.1; JOINED.
DR EMBL; AJ288506; CAC36101.1; JOINED.
DR EMBL; AJ288507; CAC36101.1; JOINED.
DR EMBL; AJ288508; CAC36101.1; JOINED.
DR EMBL; AJ288509; CAC36101.1; JOINED.
DR EMBL; AJ288510; CAC36101.1; JOINED.
DR EMBL; AJ288511; CAC36101.1; JOINED.
DR EMBL; AJ288512; CAC36101.1; JOINED.
DR EMBL; AJ288513; CAC36101.1; JOINED.
DR EMBL; AJ288514; CAC36101.1; JOINED.
DR EMBL; AJ288515; CAC36101.1; JOINED.
DR EMBL; AJ288516; CAC36101.1; JOINED.
DR EMBL; AJ288517; CAC36101.1; JOINED.
DR EMBL; AJ288518; CAC36101.1; JOINED.
DR EMBL; AJ288519; CAC36101.1; JOINED.
DR EMBL; AJ288520; CAC36101.1; JOINED.
DR EMBL; AJ288521; CAC36101.1; JOINED.
DR EMBL; AJ288522; CAC36101.1; JOINED.
DR EMBL; AJ288523; CAC36101.1; JOINED.
DR EMBL; AJ288524; CAC36101.1; JOINED.
DR EMBL; AJ288525; CAC36101.1; JOINED.
DR EMBL; AJ288526; CAC36101.1; JOINED.
DR EMBL; AJ288527; CAC36101.1; JOINED.
DR EMBL; AJ288528; CAC36101.1; JOINED.
DR EMBL; AJ288529; CAC36101.1; JOINED.
DR EMBL; AJ288530; CAC36101.1; JOINED.
DR EMBL; AJ288531; CAC36101.1; JOINED.
DR EMBL; AJ288532; CAC36101.1; JOINED.
DR EMBL; AJ288533; CAC36101.1; JOINED.
DR EMBL; AJ288534; CAC36101.1; JOINED.
DR EMBL; AJ288535; CAC36101.1; JOINED.

DR EMBL: AJ288536; CAC36101.1; JOINED.
 DR EMBL: AJ288537; CAC36101.1; JOINED.
 DR EMBL: AJ288538; CAC36101.1; JOINED.
 DR EMBL: M92993; AAA21610.1; -.
 DR EMBL: M81379; AAB19637.1; -.
 DR EMBL: M81379; AAA51556.1; -.
 DR EMBL: L08650; AAA52044.1; -.
 DR EMBL: U02519; AAA18942.1; -.

Query Match 4.5%; Score 11; DB 1; Length 1670;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dr 71 PGLPGRKDPG 81
 Db 1300 PGLPGRKDPG 1310

RESULT 9
 CA17_HUMAN STANDARD; PRT: 2944 AA.
 ID 002386; Q14054; Q16507;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Collagen alpha 1(VII) chain precursor (long-chain collagen) (LC
 DE collagen).
 GN COL7A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94327588; PubMed=8051117;
 RA Cristiano A.M., Greenspan D.S., Lee S., Uitto J.;
 RT "Cloning of human type VII collagen. Complete primary sequence of the
 RT alpha 1(VII) chain and identification of intragenic polymorphisms.";
 RL J. Biol. Chem. 269:20256-20262(1994).
 RN [2]
 RP SEQUENCE OF 128-1493 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=9338437; PubMed=1307247;
 RA Cristiano A.M., Rosenbaum L.M., Chung-Honet L.C., Parente M.G.,
 RA Woodley D.T., Pan T.C., Zhang R.Z., Chu M.-L., Burgeson R.E.,
 RA Uitto J.;
 RT "The large non-collagenous domain (NC-1) of type VII collagen is
 RT amino-terminal and chimeric. Homology to catilage matrix protein,
 RT the type III domains of fibronectin and the A domains of von
 RT Willebrand factor.";
 RL Hum. Mol. Genet. 1:475-481(1992).
 RN [3]
 RP SEQUENCE OF 815-1439 FROM N.A.
 RX MEDLINE=9334380; PubMed=1871109;
 RA Parente M.G., Chung L.C., Rymaenen J., Woodley D.T., Wynn K.W.,
 RA Bauer E.A., Mattei M.-G., Chu M.-L., Uitto J.;
 RT "Human type VII collagen: cDNA cloning and chromosomal mapping of the
 RT gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:6931-6935(1991).
 RN [4]
 RP SEQUENCE OF 369-1255 FROM N.A.
 RX MEDLINE=93107742; PubMed=1469284;
 RA Gammon W.R., Abernethy M.L., Fadilla K.M., Priayanh P.S.,
 RA Cook M.E., Wright J., Briggman R.A., Hunt S.W. II;
 RT "Noncollagenous (NC1) domain of collagen VII resembles multidomain
 RT adhesion proteins involved in tissue-specific organization of
 RT extracellular matrix";
 RL J. Invest. Dermatol. 99:691-696(1992).
 RN [5]
 RP SEQUENCE OF 340-675 FROM N.A.
 RX MEDLINE=92231902; PubMed=1567409;
 RA Tanaka T., Takahashi K., Furukawa F., Imamura S.;
 RT "Molecular cloning and characterization of type VII collagen cDNA.";

RL Biochem. Biophys. Res. Commun. 183:958-963(1992).
 RN [6]
 RP SEQUENCE OF 2395-2944 FROM N.A.
 RX MEDLINE=93271985; PubMed=8499916;
 RA Greenspan D.S.;
 RT "The carboxyl-terminal half of type VII collagen, including the non-
 RT collagenous NC-2 domain and intron/exon organization of the
 RT corresponding region of the COL7A1 gene";
 RL Hum. Mol. Genet. 2:273-278(1993).
 RN [7]
 RP SEQUENCE OF 1-87 FROM N.A.
 RX TISSUE=Placenta;
 RX MEDLINE=94375010; PubMed=8088784;
 RA Cristiano A.M., Hoffman G.G., Chung-Honet L.C., Lee S., Cheng W.,
 RA Uitto J., Greenspan D.S.;
 RT "Structural organization of the human type VII collagen gene (COL7A1),
 RT composed of more exons than any previously characterized gene";
 RL Genomics 21:169-179(1994).
 RN [8]
 RP REVIEW ON DER VARIANTS.
 RX MEDLINE=98041696; PubMed=9375848;
 RA Jaervikallio A., Pulkkinen L., Uitto J.;
 RT "Molecular basis of dystrophic epidermolysis bullosa: mutations in
 RT the type VII collagen gene (COL7A1).";
 RL Hum. Mutat. 10:338-347(1997).
 RN [9]
 RP VARIANT RDEB LYS-2798.
 RX MEDLINE=93291877; PubMed=8513326;
 RA Cristiano A.M., Greenspan D.S., Hoffman G.G., Zhang X., Tamai Y.,
 RA Lin A.H., Dietz H.C., Hovnanian A., Uitto J.;
 RT "A missense mutation in type VII collagen in two affected siblings
 RT with recessive dystrophic epidermolysis bullosa";
 RL Nat. Genet. 4:62-66(1993).
 RN [10]
 RP VARIANT DDEB SER-2040.
 RX MEDLINE=94224777; PubMed=8170945;
 RA Cristiano A.M., Rymaenen M., Uitto J.;
 RT "Dominant dystrophic epidermolysis bullosa: identification of a
 RT Gly--Ser substitution in the triple-helical domain of type VII
 RT collagen";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3549-3553(1994).
 RN [11]
 RP VARIANT PEB-DDEB CYS-2623.
 RX MEDLINE=96081220; PubMed=8541842;
 RA Cristiano A.M., Lee J.Y.-Y., Chen W.J., Laforgia S., Uitto J.;
 RT "Pretibial epidermolysis bullosa: genetic linkage to COL7A1 and
 RT identification of a glycine-to-cysteine substitution in the triple-
 RT helical domain of type VII collagen";
 RL Hum. Mol. Genet. 4:1579-1583(1995).
 RN [12]
 RP VARIANT DDEB ARG-2043.
 RX MEDLINE=95164985; PubMed=7861014;
 RA Cristiano A.M., Morricone A., Paradisi M., Angelo C., Mazzanti C.,
 RA Cavallieri R., Uitto J.;
 RT "A glycine-to-arginine substitution in the triple-helical domain of
 RT type VII collagen in a family with dominant dystrophic epidermolysis
 RT bullosa";
 RL J. Invest. Dermatol. 104:438-440(1995).
 RN [13]
 RP VARIANTS RDEB AND DDEB.
 RX MEDLINE=96220218; PubMed=8644729;
 RA Cristiano A.M., McGrath J.A., Tan K.C., Uitto J.;
 RT "Glycine substitutions in the triple-helical region of type VII
 RT collagen result in a spectrum of dystrophic epidermolysis bullosa
 RT phenotypes and patterns of inheritance";
 RL Am. J. Hum. Genet. 58:671-681(1996).
 RN [14]
 RP VARIANT RDEB ARG-2575.
 RX MEDLINE=96154068; PubMed=8592061;
 RA Shimizu H., McGrath J.A., Cristiano A.M., Nishikawa T., Uitto J.;
 RT "Molecular basis of recessive dystrophic epidermolysis bullosa:
 RT genotype/phenotype correlation in a case of moderate clinical
 RT severity.";

RL J. Invest. Dermatol. 106:119-124(1996).
 RN [15]
 RP VARIANT RDEB ARG-1782.
 RX MEDLINE=96183562; PubMed=8618018;
 RA Christiano A.M., McGrath J.A., Uitto J.;
 RT "Influence of the second COL7A1 mutation in determining the
 RT phenotypic severity of recessive dystrophic epidermolysis bullosa";
 RL J. Invest. Dermatol. 106:766-770(1996).
 RN [16]
 RP VARIANT RDEB ASP-2073.
 RX MEDLINE=96310789; PubMed=8757758;
 RA Dunnill M.G.S., McGrath J.A., Richards A.J., Christiano A.M.,
 RA Uitto J., Pope F.M., Eady R.A.J.;
 RT "Clinicopathological correlations of compound heterozygous COL7A1
 RT mutations in recessive dystrophic epidermolysis bullosa";
 RL J. Invest. Dermatol. 107:171-177(1996).
 RN [17]
 RP VARIANTS RDEB W-1982; G-2008; A-2025; E-2049; G-2063; W-2063 AND
 RP R-2575.
 RX MEDLINE=97465605; PubMed=9326325;
 RA Hovnanian A., Rochat A., Bodemer C., Petit E., Rivers C.A., Prost C.,
 RA Prats Y., Christiano A.M., Uitto J., Lathrop M., Barrandon Y.;
 RA de Prost Y.;
 RT "Characterization of 18 new mutations in COL7A1 in recessive
 RT dystrophic epidermolysis bullosa provides evidence for distinct
 RT molecular mechanisms underlying defective anchoring fibril
 RT formation";
 RL Am. J. Hum. Genet. 61:599-610(1997).
 RN [18]
 RP VARIANT RDEB ARG-1652.
 RX MEDLINE=98106792; PubMed=9444387;
 RA Csechalmi-Friedman P.B., Karpatis S., Horvath A., Christiano A.M.;
 RT "Identification of a glycine substitution and a splice site mutation
 RT in the type VII collagen gene in a proband with mild recessive
 RT dystrophic epidermolysis bullosa";
 RL Arch. Dermatol. Res. 289:640-645(1997).
 RN [19]
 RP VARIANTS DEB ARG-2009 AND ARG-2043.
 RX MEDLINE=97358588; PubMed=9215684;
 RA Winderig J.-O., Hammant-Hausali N., Nilsen O., Anton-Lamprecht I.,
 RA Naylor S.L., Kerbacher K., Zimmermann M., Krajci P.,
 RA Gede-Dahl T. Jr., Bruckner-Tuderman L.;
 RT "Modulation of disease severity of dystrophic epidermolysis bullosa by
 RT a splice site mutation in combination with a missense mutation in the
 RT COL7A1 gene";
 RL Hum. Mol. Genet. 6:1125-1135(1997).
 RN [20]
 RP VARIANTS DEB ASP-1519; ASP-2006; GLU-2015 AND ARG-2034.
 RX MEDLINE=98334662; PubMed=9668111;
 RA Hammant-Hausali N., Schumann H., Raghunath M., Kilgus O., Luetthi U.,
 RA Luger T., Bruckner-Tuderman L.;
 RT "Some, but not all, glycine substitution mutations in COL7A1 result in
 RT intracellular accumulation of collagen VII, loss of anchoring
 RT fibrils, and skin blistering";
 RL J. Biol. Chem. 273:19228-19234(1998).
 RN [21]
 RP VARIANTS DEB CYS-2008; ARG-2207 AND SER-2775.
 RX MEDLINE=98410969; PubMed=9740253;
 RA Kon A., Pulkkinen L., Ishida-Yamamoto A., Hashimoto I., Uitto J.;
 RT "Novel COL7A1 mutations in dystrophic forms of epidermolysis
 RT bullosa";
 RL J. Invest. Dermatol. 111:534-537(1998).
 RN [22]
 RP VARIANT RDEB ARG-1347.
 RX MEDLINE=99019477; PubMed=9804332;
 RA Terracina M., Posteraro P., Schubert M., Sonego G., Atzori F.,
 RA Zambuno G., Bruckner-Tuderman L., Castiglia D.;
 RT "Compound heterozygosity for a recessive glycine substitution and a
 RT splice site mutation in the COL7A1 gene causes an unusually mild form
 RT of localized recessive dystrophic epidermolysis bullosa";
 RL J. Invest. Dermatol. 111:744-750(1998).
 RN [23]
 RP VARIANTS DEB TRP-2034; VAL-2040; ARG-2043; ARG-2064 AND ASP-2713.

RX MEDLINE=99072663; PubMed=9856843;
 Query Match 4.5%; Score 11; DB 1; Length 2944;
 Best Local Similarity 100.0%; Pred. No. 0.38;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 71 GAGPGRGPG 81
 DB 1328 PGLPGRGPG 1338
 RESULT 10
 CAC2_HAECC STANDARD; PRT; 210 AA.
 ID CAC2_HAECC
 AC P16252;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cuticle collagen 2C (Fragment).
 GN 2C.
 OS Haemonchus contortus (Barber pole worm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Trichostrongylidae; Haemonchidae; Haemonchinae; Haemonchus.
 OX NCBI_TaxID=6289;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90136718; PubMed=2615789;
 RA Sharnansky L.M., Pratt D., Boissvenne R.J., Cox G.N.;
 RT "Cuticle collagen genes of Haemonchus contortus and Caenorhabditis
 RT elegans are highly conserved";
 RL Mol. Biochem. Parasitol. 37:73-86(1989).
 CC -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
 CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
 CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
 CC -1- MISCELLANEOUS: THIS PROTEIN SHOWS A POTENTIAL TRIPLE-HELICAL
 CC REGIONS, WHICH CONTAIN GLYCINE AS EVERY THIRD AMINO ACID.
 CC -1- MISCELLANEOUS: IN ALL NEMATODE CUTICLE COLLAGENS, THE POLYPEPTIDE
 CC CHAINS ARE COMPLEXED WITHIN THE CUTICLE BY DISULFIDE BONDS AND
 CC OTHER TYPES OF COVALENT CROSS-LINKS.
 CC -1- SIMILARITY: BELONGS TO THE CUTICULAR COLLAGEN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; J04670; AAA29172.1; -;
 CC DR InterPro; IPR000087; Collagen.
 CC Pfam; PF01391; Collagen; 3.
 CC Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
 CC NON_TER 1
 CC SEQUENCE 210 AA; 19562 MW; E15FA9A2DF3D74B CRC64;
 Query Match 4.1%; Score 10; DB 1; Length 210;
 Best Local Similarity 100.0%; Pred. No. 0.31;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 GAGPAGGEGK 66
 DB 151 GAGPAGGEGK 160
 RESULT 11
 PSPA_CAVPO STANDARD; PRT; 247 AA.
 ID PSPA_CAVPO
 AC P50403;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)

Db 39 GLPGRDRG 48

|||||

RESULT 13
PSPA CANFA STANDARD; PRT; 248 AA.
ID PSPA CANFA
AC P06908:
DT 01-JAN-1998 (Rel. 06, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A) (PSPAP).
GN SFTPA1 OR SFTPA OR SFTPL.
OS Caris familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
[1]
RP SEQUENCE FROM N.A., AND SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=86016705; PubMed=3863100;
RA Benson B., Hawgood S., Schilling J., Clements J., Damm D., Cordell B., White R.T.;
RT "Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino acid sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 82:6379-6383(1985).
RN [2]
RP C-TYPE LECTIN DOMAIN.
RX MEDLINE=87115834; PubMed=3808053;
RA Patchy L.;
RT "Is lung surfactant protein a lectin-collagen hybrid?";
RL Nature 325:490-490(1987).
CC -1- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS ESSENTIAL FOR NORMAL RESPIRATION.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10% PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS, CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11769; AAA30887.1; -.
DR PIR; A25296; LNDGPS.
DR HSSP; P22897; 1EG3.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C; 1.
DR Pfam; PF01391; COLLAGEN; 2.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS00641; C-TYPE LECTIN_2; 1.
KM Glycoprotein; Calcium, Surface film; Gaseous exchange; Hydroxylation; Signal; Lectin; Collagen; Repeat.
FT SIGNAL 1 17
FT CHAIN 18 248
FT DOMAIN 26 100
FT FT 153 248
FT DISULFID 155 246
FT DISULFID 224 238
FT CARBOHYD 20 20
A. PULMONARY SURFACTANT-ASSOCIATED PROTEIN
COLLAGEN-LIKE.
C-TYPE LECTIN (SHORT FORM).
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 207 207 N-LINKED (GLCNAC. .) (PROBABLE).
SQ SEQUENCE 248 AA; 26268 MM; 340F95D4E2502C0 CRC64;

Query Match 4.1%; Score 10; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 45 GLPGRDRG 54
Db 40 GLPGRDRG 49

RESULT 14
PSPA HUMAN STANDARD; PRT; 248 AA.
ID PSPA HUMAN
AC P07714:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A) (PSPAP) (Alveolar proteinosis protein) (35 kDa pulmonary surfactant-associated protein).
GN SFTPA1 OR SFTPA OR SFTPL OR PSAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86250832; PubMed=3755136;
RA Flores J., Steinbrink R., Jacobs K., Phelps D., Kriz R., Recny M., Sultzman L., Jones S., Taesch H.W., Frank H.A., Fritsch E.F.;
RT "Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfactant-associated protein."
RL J. Biol. Chem. 261:9029-9033(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86013366; PubMed=2995821;
RA White R.T., Damm D., Miller J., Spratt K., Schilling J., Hawgood S., Benson B., Cordell B.;
RT "Isolation and characterization of the human pulmonary surfactant apoprotein gene."
RL Nature 317:361-363(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92198680; PubMed=1372511;
RA Katyal S.L., Singh G., Locker J.L.;
RT "Characterization of a second human pulmonary surfactant-associated protein SP-A gene."
RL Am. J. Respir. Cell Mol. Biol. 6:446-452(1992).
CC -1- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS ESSENTIAL FOR NORMAL RESPIRATION.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10% PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS, CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13686; AAA60211.1; -.
DR EMBL; K03475; AAA36520.1; -.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:31:09 ; Search time 31 Seconds

(without alignments)
1615.144 Million cell updates/sec

Title: US-09-944-403-42

Perfect score: 243
Sequence: 1 MRPLVLILLGLAAGSPPLD.....DSTFGFLVSDMHSSPVFA 243

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	25.5	243	11 Q8R002	Q8R002 mus musculus
2	11	4.5	27	4 Q9UCJ3	Q9UCJ3 homo sapien
3	11	4.5	289	5 Q19813	Q19813 caenorhabdi
4	11	4.5	306	2 Q9ACN2	Q9ACN2 streptococc
5	11	4.5	375	4 Q8TCD8	Q8TCD8 homo sapien
6	11	4.5	378	4 Q8N1X4	Q8N1X4 sus scrofa
7	10	4.1	57	6 Q02842	Q02842 papio cynoc
8	10	4.1	238	13 Q57451	Q57451 gallus galli
9	10	4.1	246	13 Q91907	Q91907 carassius a
10	10	4.1	248	6 Q9T06	Q9T06 ovis aries
11	10	4.1	248	6 Q9N0G1	Q9N0G1 equus caball
12	10	4.1	248	6 Q951A8	Q951A8 equus caball
13	10	4.1	248	11 Q9CQ11	Q9CQ11 mus musculus
14	10	4.1	254	13 Q987B4	Q987B4 gallus galli
15	10	4.1	256	13 Q91909	Q91909 cyprinus ca
16	10	4.1	282	5 Q8WP36	Q8WP36 suberites d

17	10	4.1	358	12 Q9J3U4	Q9J3U4 ectocarpus
18	10	4.1	447	12 Q9TYJ3	Q9TYJ3 ateline her
19	10	4.1	694	16 Q977A5	Q977A5 streptococ
20	10	4.1	812	5 Q06452	Q06452 ephyacia m
21	10	4.1	1761	5 Q18407	Q18407 drosophila
22	10	4.1	1940	5 Q9VWV5	Q9VWV5 drosophila
23	9	3.7	182	11 Q9CYS4	Q9CYS4 mus musculu
24	9	3.7	260	4 Q9UGS2	Q9UGS2 homo sapien
25	9	3.7	260	11 Q9OXU9	Q9OXU9 ratu
26	9	3.7	461	13 Q90Y19	Q90Y19 brachydanio
27	9	3.7	567	10 Q9SMQ3	Q9SMQ3 artemisia a
28	9	3.7	589	16 Q9AK64	Q9AK64 streptomyce
29	9	3.7	684	5 P90679	P90679 arenicola m
30	9	3.7	695	11 Q64348	Q64348 mus musculu
31	9	3.7	751	11 Q60709	Q60709 mus musculu
32	9	3.7	763	11 Q61482	Q61482 mus musculu
33	9	3.7	1752	5 Q07265	Q07265 strongyloce
34	9	3.7	2288	5 Q23081	Q23081 caenorhabdi
35	9	3.7	2315	5 Q952K3	Q952K3 caenorhabdi
36	8	3.3	46	11 Q63074	Q63074 rattus norv
37	8	3.3	52	16 Q9H209	Q9H209 pseudomonas
38	8	3.3	87	2 Q45313	Q45313 bacillus me
39	8	3.3	107	4 Q9BQ77	Q9BQ77 homo sapien
40	8	3.3	109	11 Q9CVJ2	Q9CVJ2 mus musculu
41	8	3.3	111	16 Q9RDP9	Q9RDP9 streptomyce
42	8	3.3	149	16 Q9JXB1	Q9JXB1 neisseria m
43	8	3.3	149	16 Q9JWV2	Q9JWV2 neisseria m
44	8	3.3	164	16 P74713	P74713 synchocyst
45	8	3.3	170	6 Q9GLV5	Q9GLV5 macaca mulla

ALIGNMENTS

RESULT 1

Q8R002 ID Q8R002 PRELIMINARY; PRT; 243 AA.

AC Q8R002; 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Similar to DKFZ586B0621 protein (hypothetical 25.4 kDa protein).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.; Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Strausberg R.; Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC025174; AAH25174.1; -

DR EMBL; BC025174; AAH25174.1; -

KW Hypothetical protein.

SQ SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64;

Query Match 182 FGGMPKRSASLGAMVRLBPEDQVWVGVDYIGIYASIKTSTFGFLVSDMHSSPV 241

Best Local Similarity 100.0%; Pred. No. 4, 8e-52; Length 243;

Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 FGGMPKRSASLGAMVRLBPEDQVWVGVDYIGIYASIKTSTFGFLVSDMHSSPV 241

Db 182 FGGMPKRSASLGAMVRLBPEDQVWVGVDYIGIYASIKTSTFGFLVSDMHSSPV 241

QY 242 FA 243

Db 242 FA 243

RESULT 2

09UCJ3
ID 09UCJ3 PRELIMINARY; PRT; 27 AA.
AC 09UCJ3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Surfactant protein D (Fragment).
OS Homo sapiens (Human).
OC Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE.
RX MEDLINE=93142849; PubMed=8424457;
RA Crouch E., Persson A., Chang D.;
RT "Accumulation of surfactant protein D in human pulmonary alveolar
RT proteinosis.";
RL Am. J. Pathol. 142:241-248(1993).
DR InterPro: IPR000087; Collagen.
SQ SEQUENCE 27 AA; 2696 MW; 8F770650E2A967CA CRC64;
Query Match 4.5%; Score 11; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 GLPGRDGRDR 55
DB 1 GLPGRDGRDR 11
RESULT 3
019813 PRELIMINARY; PRT; 289 AA.
ID 019813;
AC 019813;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F26F12.1 protein.
CN F26F12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RT Science 287:2012-2018(1998).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Wilson R., Bentley D., Gattung S.;
RA "The sequence of C. elegans cosmid F26F12.";
RT Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Waterston R.;
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; U55373; AAC25888.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR002486; Col. cuticle_N.
DR Pfam; PF01391; Collagen; 3.
DR Pfam; PF01484; Col. cuticle_N; 1.
SQ SEQUENCE 289 AA; 28740 MW; DFFB78B2B34B29DF CRC64;
Query Match 4.5%; Score 11; DB 5; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GPRGEAGPAG 90
DB 222 GPRGEAGPAG 232
RESULT 4
09ACN2 PRELIMINARY; PRT; 306 AA.
ID 09ACN2;
AC 09ACN2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Sc1B protein.
GN SC1B.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillae; Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxId=1314;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=655;
RX MEDLINE=21097281; PubMed=11158359;
RA Whatmore A.M.;
RT "Streptococcus pyogenes sc1B encodes a putative hypervariable surface
RT protein with a collagen-like repetitive structure.";
RL Microbiology 147:419-429(2001).
DR EMBL; AJ301809; CAC33778.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam; PF01391; Collagen; 1.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; lpxTG anchor; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KM Transmembrane.
SQ SEQUENCE 306 AA; 31833 MW; 57EBDC244DF17600 CRC64;
Query Match 4.5%; Score 11; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 GPRGEAGPAG 91
DB 148 GPRGEAGPAG 158
RESULT 5
08TCD8 PRELIMINARY; PRT; 375 AA.
ID 08TCD8;
AC 08TCD8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 37.7 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC022318; AAH22318.1; -.
KM Hypothetical protein.
SQ SEQUENCE 375 AA; 37655 MW; CCB7375D3C86421A CRC64;
Query Match 4.5%; Score 11; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 GLPGRDGRDR 55
|||||

Db 46 GLPGRDGRDGR 56

RESULT 6

Q9N1X4 PRELIMINARY; PRT; 378 AA.

AC Q9N1X4; 09N1X4;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Lung surfactant protein D precursor.

GN SFTPM.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20109098; PubMed=10640760;

RA van Eijk M., Haagsman H.P., Skinner T., Archibald A.L., Reid K.B.M., Lawson P.R.;

RT "Porcine Lung Surfactant Protein D (SP-D): cDNA cloning, chromosomal localization and tissue distribution."

RT J. Immunol. 164:1442-1450(2000).

RL EMBL; AF132496; AAF22145.2; -.

DR HSP; P35247; 1B08.

DR Interpro; IPR000087; Collagen.

DR Interpro; IPR001304; Lectin_C.

DR Pfam; PF01391; Collagen; 3.

DR Pfam; PF00059; lectin_c; 1.

DR Prodom; PD000007; Collagen; 2.

DR SMART; SM00034; CLECT; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

KM SIGNAL.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 378 LUNG SURFACTANT PROTEIN D.

SQ SEQUENCE 378 AA; 37986 MW; 35048C1E56C34ID CRC64;

Query Match

Best Local Similarity 4.5%; Score 11; DB 6; Length 378;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDGR 55

Db 46 GLPGRDGRDGR 56

RESULT 7

002842 PRELIMINARY; PRT; 57 AA.

ID 002842

AC 002842;

DT 01-JUN-1997 (TREMBlrel. 04, Created)

DT 01-JUN-1997 (TREMBlrel. 04, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Surfactant protein A2 (fragment).

OS Papio cynocephalus (Yellow baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopitheciinae; Papio.

OX NCBI_TaxID=9556;

OK [1]

RN SEQUENCE FROM N.A.

RP TISSUE=KIDNEY;

RX MEDLINE=97053398; PubMed=8897910;

RA Gao E., Wang Y., McCormick S.M., Li J., Seidner S.R., Mendelson C.R.;

RT "Characterization of two baboon surfactant protein A genes."

RL Am. J. Physiol. 271:L617-L630(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RA Gao E., Wang Y., McCormick S.M., Li J., Seidner S.R., Mendelson C.R.;

RL Submitted (May-1997) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF003897; AAB61295.1; -.

DR EMBL; AF003896; AAB61294.1; -.

DR Interpro; IPR000087; Collagen.

DR Prodom; PD000007; Collagen; 1.

FT NON TER 57

SQ SEQUENCE 57 AA; 5584 MW; B25149EC822F643C CRC64;

Query Match

Best Local Similarity 4.1%; Score 10; DB 6; Length 57;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDGR 54

Db 40 GLPGRDGRDGR 49

RESULT 8

057451 PRELIMINARY; PRT; 238 AA.

ID 057451

AC 057451;

DT 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Mannan-binding lectin (Fragment).

GN CMBL.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=WHITE LEGHORN; TISSUE=LIVER;

RA Laursen S.B.;

RL EMBL; AF022226; AAB94071.1; -.

DR HSP; P19999; 1YTT.

DR Interpro; IPR000087; Collagen.

DR Interpro; IPR001304; Lectin_C.

DR Pfam; PF01391; Collagen; 1.

DR Pfam; PF00059; lectin_c; 1.

DR Prodom; PD000007; Collagen; 1.

DR SMART; SM00034; CLECT; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

KM LECTIN.

FT NON TER 238

SQ SEQUENCE 238 AA; 25645 MW; B5C9B5197AAB64E3 CRC64;

Query Match

Best Local Similarity 4.1%; Score 10; DB 13; Length 238;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDGR 54

Db 31 GLPGRDGRDGR 40

RESULT 9

091907 PRELIMINARY; PRT; 246 AA.

ID 091907

AC 091907;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Mannose binding-like lectin precursor (fragment).

GN MBL.

OS Carassius auratus (Goldfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthys; Cypriniformes; Cyprinidae; Carassius.

OX NCBI_TaxID=7957;

RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=20456722; PubMed=11003389;
 RA Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjodt K.;
 RT "The homologue of mannose-binding lectin in the carp family Cyprinidae
 is expressed at high level in spleen, and the deduced primary
 structure predicts affinity for galactose.";
 RT Immunogenetics 51:955-964(2000).
 RL EMBL: AF227739; AAF63470.1; -.
 DR HSSP: P35247; 1B08.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF01391; Collagen; 2.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 KM Lectin; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 13
 FT VARIANT 145 145 S -> P.
 SQ SEQUENCE 246 AA; 25709 MW; AB692282D289DD5 CRC64;
 Query Match 4.1%; Score 10; DB 13; Length 246;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 45 GLPGRDGRDG 54
 DB 35 GLPGRDGRDG 44
 RESULT 10
 C9T06 PRELIMINARY; PRT; 248 AA.
 ID C9T06
 AC O9T06; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pulmonary surfactant protein A (Pulmonary surfactant-associated
 protein A).
 DE SPAS OR SP-A.
 GN Ovis aries (Sheep).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20215262; PubMed=10749753;
 RA Brems G.A., Yao L.-J., Inchley K., Brickenden A., Han V.K.M.,
 RA Gioia A., Challis J.R.G., Posemayer P.;
 RT "Ovine surfactant protein CDNA: use in studies on fetal lung growth
 and maturation after prolonged hypoxemia.";
 RT Am. J. Physiol. 278:L754-L764(2000).
 RL Am. J. Physiol. 278:L765-L778(2000).
 DR EMBL: AF211856; AAF18995.1; -.
 DR HSSP: P22897; 1BGS.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF01391; Collagen; 2.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 SQ SEQUENCE 248 AA; 26047 MW; B71133E005C9A5C1 CRC64;

SQ SEQUENCE 248 AA; 26394 MW; D65E7293BBF1FD9 CRC64;
 Query Match 4.1%; Score 10; DB 6; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 45 GLPGRDGRDG 54
 DB 40 GLPGRDGRDG 49
 RESULT 11
 C9N0G1 PRELIMINARY; PRT; 248 AA.
 ID C9N0G1
 AC O9N0G1; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pulmonary surfactant protein A.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=THOROUGHRED; TISSUE=LUNG;
 RA Hobo S.;
 RT "Molecular cloning of equine pulmonary surfactant proteins.";
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB015963; BAA97976.1; -.
 DR HSSP: P35247; 1B08.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF01391; Collagen; 2.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 SQ SEQUENCE 248 AA; 26000 MW; BBE12EFD05C2B8D1 CRC64;
 Query Match 4.1%; Score 10; DB 6; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 45 GLPGRDGRDG 54
 DB 40 GLPGRDGRDG 49
 RESULT 12
 C9S1B8 PRELIMINARY; PRT; 248 AA.
 ID C9S1B8
 AC O9S1B8; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pulmonary surfactant-associated protein A.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Weber B.L., Hospes R., Gortner L.;
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF400580; AAL07690.1; -.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF01391; Collagen; 2.
 DR Pfam: PF00059; lectin_c; 1.
 DR PROSITE: PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 SQ SEQUENCE 248 AA; 26047 MW; B71133E005C9A5C1 CRC64;

Query Match 4.1%; Score 10; DB 6; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
 Db 40 GLPGRDGRDG 49

RESULT 13
 Q9C011 PRELIMINARY; PRT; 248 AA.
 AC Q9C011; 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Surfactant associated protein A.
 GN SPTPA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRIO, AND LUNG;
 RX MEDLINE=2108566Q; PubMed=1121851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arkawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alkawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda M., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staudt F., Suzuki K., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guetlinch S., Hill D., Hofmann M., Hume D.A., Kamlay M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlschki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse CDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK011333; BAB27551.1; -
 DR EMBL; AK004620; BAB23416.1; -
 DR EMBL; AK004788; BAB23565.1; -
 DR HSSP; P35247; 1B08
 DR MGD; MGI:109518; Stpa.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF01391; Collagen; 2.
 DR Pfam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1;
 DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1;
 DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1;
 SQ SEQUENCE 248 AA; 26183 MW; 8A5670CFAD3EB9B6 CRC64;

Query Match 4.1%; Score 10; DB 11; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
 Db 40 GLPGRDGRDG 49

RESULT 14
 Q98TA4 PRELIMINARY; PRT; 254 AA.

Q98TA4;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Mannose-binding lectin protein precursor.
 OS Gallus gallus (Chicken)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20456722; PubMed=11003389;
 RA Vltved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjold K.;
 RT "The homologue of mannose-binding lectin in the carp family Cyprinidae
 is expressed at high level in spleen, and the deduced primary
 structure predicts affinity for galactose."
 RL Immunogenetics 51:955-964(2000).
 DR EMBL; AF231714; AAK30298.1; -
 DR HSSP; P19999; 1YTT.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR000005; HTHAraC.
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF01391; Collagen; 1.
 DR Pfam; PF00059; Lectin_c; 1.
 DR ProDom; PD000007; Collagen; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
 DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
 KW Lectin; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 254
 FT VARIANT 234 254 D -> V.
 SQ SEQUENCE 254 AA; 27376 MW; C924428643441MED CRC64;

Query Match 4.1%; Score 10; DB 13; Length 254;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
 Db 47 GLPGRDGRDG 56

RESULT 15
 Q91909 PRELIMINARY; PRT; 256 AA.
 AC Q91909;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Mannose binding-like lectin precursor.
 GN MBL.
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OC NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=20456722; PubMed=11003389;
 RA Vltved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjold K.;
 RT "The homologue of mannose-binding lectin in the carp family Cyprinidae
 is expressed at high level in spleen, and the deduced primary
 structure predicts affinity for galactose."
 RL Immunogenetics 51:955-964(2000).
 DR EMBL; AF227737; AAF63468.1; -
 DR HSSP; P35247; 1B08.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF01391; Collagen; 2.

DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_2; 1.
KW lectin; signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 256 MANNOSE BINDING-LIKE LECTIN.
FT VARIANT 235 235 S -> T.
SQ SEQUENCE 256 AA; 26934 MW; D019291D116730D CRC64;

Query Match 4.1%; Score 10; DB 13; Length 256;
Best Local Similarity 100.0%; Pred.No. 0.22;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
|||
Db 45 GLPGRDGRDG 54

Search completed: June 20, 2003, 11:42:03
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 08:43:49 ; Search time 62 Seconds
(without alignments)
522.256 Million cell updates/sec

Title: US-09-944-403-42
Perfect score: 1325
Sequence: 1 MRPLVLLGLAAGSPPLD.....DSTFGSLVYSWMHSEVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDS2/gcgdata/geneseq/emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1325	100.0	243	20	AAV06481 Human tumour-associ
2	1325	100.0	243	20	AAV17827 Human PRO344 prote
3	1325	100.0	243	20	AAW97984 Human adipocyte-sp
4	1325	100.0	243	21	AAAB3461 Human PRO344 prote
5	1325	100.0	243	21	AAV71468 Human PRO344 prote
6	1325	100.0	243	21	AAV93688 Amino acid sequenc
7	1325	100.0	243	21	AAAB01318 Human PRO344 polyp
8	1325	100.0	243	22	AAU12352 Human PRO344 polyp
9	1325	100.0	243	22	AAAB5815 Human TANGO 253 SE
10	1325	100.0	243	22	AAAB9593 Human adipocyte co

11	1325	100.0	243	22	AAAB9599 Human adipocyte co
12	1323	99.8	243	22	AAAB5891 Human secreted pro
13	1321	99.7	243	22	AAAB5888 Human secreted pro
14	1321	99.7	243	22	AAAB5889 Human secreted pro
15	1321	99.7	243	22	AAAB5890 Human secreted pro
16	1271	95.9	243	21	AAV76040 Rat skin cell prote
17	1271	95.9	243	22	AAAB5979 Skin cell protein,
18	1271	95.9	243	23	AAAB72179 Rat protein isolat
19	1258	94.9	243	22	AAAB5820 Murine TANGO 253 S
20	1256	94.8	243	22	AAAB5816 Human mature TANG
21	1256	94.8	243	22	AAAB5899 Murine secreted pr
22	1254	94.6	243	22	AAAB5897 Murine secreted pr
23	1254	94.6	243	22	AAAB5898 Murine secreted pr
24	1243.5	93.8	242	22	AAAB5896 Murine mature TANG
25	1200	90.6	228	22	AAAB5821 Novel human diagno
26	1183	89.3	220	22	AAAB12724 Human secreted pro
27	1048	79.1	201	23	AAAB21663 Human secreted pro
28	675	50.9	128	22	AAAB5819 Human TANGO 253 Cl
29	650	49.1	128	22	AAAB5824 Human TANGO 253 C
30	478	36.1	151	22	ABG12723 Novel human diagno
31	449.5	33.9	225	23	ABAB80583 Human sbg1033026C1
32	442.5	33.4	333	23	ABAB80582 Human sbg1033026C1
33	433.5	32.7	247	18	AAW09107 Mouse OBG3 protein
34	433.5	32.7	247	22	AAAB05528 Mouse OBG3 protein
35	433.5	32.7	247	23	ABAB08222 Mouse acrp30 prote
36	425	32.1	244	18	AAW09108 Mouse OBG3 protein
37	424.5	32.0	247	22	AAAB05527 Mouse OBG3 protein
38	424.5	32.0	247	23	ABAB08221 Mouse adipoc prote
39	424	32.0	244	20	AAV21807 Adipose most abund
40	424	32.0	244	21	AAAB30233 Human adipocyte co
41	424	32.0	244	21	AAV71035 Human APM1 (Adipos
42	424	32.0	244	22	AAAB05529 Human OBG3 protein
43	424	32.0	244	22	AAAB5828 Human adipocyte co
44	424	32.0	244	22	AAAB49592 Human ACRP30 prote
45	424	32.0	244	22	AAAB49598 Human ACRP30 prote

ALIGNMENTS

RESULT 1	
ID	AAV06481 standard; Protein; 243 AA.
XX	
AC	AAV06481;
XX	
DT	27-SEP-1999 (first entry)
XX	
DE	Human tumour-associated protein PRO344.
XX	
KW	PRO344; UNQ303; cancer; tumour; diagnosis; therapy; human.
XX	
OS	Homo sapiens.
XX	
XX	
FH	Key
FT	Peptide
FT	/note= "signal peptide"
FT	Protein
FT	/note= "mature protein"
FT	Modified-site
FT	/note= "N-myristoylated"
FT	Modified-site
FT	/note= "N-myristoylated"
XX	
XX	
XX	WO9935170-A2.
XX	
XX	15-JUL-1999.
PD	
XX	
XX	05-JAN-1999;
XX	
XX	99WO-US00106.
XX	
PR	20-NOV-1998;
PR	98US-0109304.
PR	05-JAN-1998;
PR	98US-0070440.
PR	29-APR-1998;
PR	98US-0083500.

ID		AAW97984 standard; Protein; 243 AA.
AC	AAW97984;	
DT	21-JUN-1999	(first entry)
DE		Human adipocyte-specific protein zsig39.
KM	Adipocyte-specific protein; zsig39; human; fatty acid metabolism; energy balance; nutrition; antimicrobial.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	Peptide	1..15 /note= "signal peptide, alternatively the signal peptide comprises residues 1..18"
FT	Protein	16..243 /note= "mature protein, alternatively the mature protein comprises residues 19..243 (specifically claimed in Claim 4)"
FT	Domain	30..96 /note= "collagen-like domain"
FT	Domain	98..243 /note= "globular domain"
FT	Region	105..109 /note= "beta strand"
FT	Region	128..130 /note= "beta strand"
FT	Region	136..139 /note= "beta strand"
FT	Region	143..146 /note= "beta strand"
FT	Region	164..171 /note= "beta strand"
FT	Region	176..182 /note= "beta strand"
FT	Region	187..200 /note= "beta strand"
FT	Region	204..210 /note= "beta strand"
FT	Region	226..231 /note= "beta strand"
FT	Domain	111..135 /note= "receptor binding domain"
FT	Domain	170..174 /note= "receptor binding domain"
PB	WO910492-A1.	
PD	04-MAR-1999.	
PP	26-AUG-1998;	98WO-US17724.
PR	26-AUG-1997;	9TUS-0056983.
PA	(ZYMO) ZYMOGENETICS INC.	
PI	Humes JM, Sheppard PO,	
DR	WPI, 1999-204665/17.	
N-PSTDB	N-PSTDB; AAX24684.	
Zsg39	Zsg39 protein - used to modulate fatty acid metabolism	
Claim 1	111-112; 132pp; English.	
This polypeptide	comprises human adipocyte-specific protein zsig39, a protein that modulates free fatty acid metabolism. zsig39 is a member of a family of proteins having a globular domain and a collagen-like domain capable of dimerisation or oligomerisation. zsig39 polypeptides were initially identified by querying an EST database for secretory signal sequences characterised by an upstream	

	CC	methionine start site, a hydrophobic region of approximately 13 amino acids and a cleavage site. A single EST sequence was discovered, and the novel polypeptide encoded by the full-length cDNA allowed the identification of a homologue relationship with adipocyte complement related protein Acyrp30 and adipocytes secreted from apM1. A full-length clone (see AAZ4684) was obtained from a lung tissue library. Expression vectors, cultured cells and a method of producing zsig99 polypeptides are claimed, as well as zsig99 polypeptides having N- or C-terminal affinity tags, toxins, radionucleotides, enzymes or fluorophores, fusion proteins including zsig99 polypeptides; an antibody that specifically binds to an epitope of zsig99, and a method for modulating free fatty acid metabolism by administering a zsig99 polypeptide. The zsig99 polypeptide may also be used in organ preservation, for cryopreservation, for surgical pretreatment to prevent injury due to ischemia and/or inflammation, and as an antimicrobial agent.
	XX	promoting lysis or phagocytosis of infectious agents.
SQ	Sequence	243 AA;
Query Match	Beat Local Similarity	100.0%; Score 1325; DB 20; Length 243;
Matches	Conservative	243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dy	I MRPLVLILLGLAAGSPFLDNKIPSLCPHRLPPTFGHHGSGCLPGDRGRDGDGAPG	60
Dd	1 MRPLLVTLLLGIAAGSPPLDNKKIPSICPHRGLPTFCHHGSGQLPRDRGRDGADG	60
Dd	61 APGGKGSGRGRLGPDPGDEBAGPACGTGPAECSTVRRARAFKSSESVPSPSD	120
Dd	61 APGGKGSGRGRLGPDPGDEBAGPACGTGPAECSTVRRARAFKSSESVPSPSD	120
Dy	121 APLEPDRLTVNEQGSHYAVTGKFTCOVGGVVFAYFAHTVVYPASIQFDLVKNGESIASPFQ	180
Dd	121 APLEPDRLTVNEQHGHYAATVKCTFCOYGCVVFYFAHVAAIVVRASIQLFDLVNGESIASPFQ	180
Dy	181 PFGGWPKRPASISGSAMRWLRPEEDVOYWQVGYDTITGIYSIKTSTSRSGLFYSDMHSSP	240
Dd	181 FFGGMWPKRPASISGSAMRWLREDDVMWGVDIIGVIYASIKTSTSRSGLFYSDMHSSP	240
Oy	241 VEA 243	
Dd	241 VEA 243	
RESULT 4	AAB33461	
ID	AAB33461 standard; Protein; 243 AA.	
AX	AAB33461;	
XT	29-JAN-2001 (first entry)	
DE	Human PRO344 protein UNQ303 SEQ ID NO:241.	
KX	Human; immune related disease; diagnosis; antiinflammatory; cardiant;	
KX	dermatological; antitarrthritis; arthrheumatic; immunosuppressive;	
KX	haemostatic; antidiabetic; nocotropic; neuroprotective;	
KX	antianemic; hepatotropic; vitruicide; antiporiastic; antiallergic;	
KX	autoasthmatic; systemic lupus erythematosus; rheumatoid arthritis;	
KX	osteoarthritis; spondylarthropathy; systemic scleriosis; sarcoidosis;	
KX	idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;	
KX	systemic vasculitis; autoimmune haemolytic anemia; diabetes mellitus;	
KX	autoimmune thrombocytopaenia; immune-mediated renal disease;	
KX	demyelinating disease; hepatoiliary disease; Whipple's disease;	
KX	inflamatory bowel disease; gluten-sensitive enteropathy;	
KX	autoimmune disease; immune-mediated skin disease; allergic disease;	
KX	immunological disease; transplantation associated disease;	
KX	graft rejection; graft-versus-host-disease.	
OS	Homo sapiens.	
XX		
PX	WO200053758-A2.	

XX 14-SEP-2000.
PD
XX
PF 02-MAR-2000; 2000MO-US05841.
XX
PR 08-MAR-1999; 99MO-US05028.
PR 10-MAR-1999; 99MO-0123618.
PR 12-MAR-1999; 99US-0123957.
PR 23-MAR-1999; 99US-0125775.
PR 12-APR-1999; 99US-0128849.
PR 20-APR-1999; 99MO-US08615.
PR 28-APR-1999; 99US-0131445.
PR 04-MAY-1999; 99US-0133371.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99MO-US12252.
PR 23-JUN-1999; 98US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0146598.
PR 28-JUL-1999; 99US-0146222.
PR 01-SEP-1999; 99MO-US20111.
PR 08-SEP-1999; 99MO-US20594.
PR 13-SEP-1999; 99MO-US20944.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
PR 05-OCT-1999; 99MO-US23089.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99MO-US28214.
PR 30-NOV-1999; 99MO-US28313.
PR 01-DEC-1999; 99MO-US28409.
PR 01-DEC-1999; 99MO-US28301.
PR 02-DEC-1999; 99MO-US28634.
PR 02-DEC-1999; 99MO-US28551.
PR 02-DEC-1999; 99MO-US28564.
PR 02-DEC-1999; 99MO-US28095.
PR 16-DEC-1999; 99MO-US30095.
PR 20-DEC-1999; 99MO-US30999.
PR 30-DEC-1999; 99MO-US31279.
PR 05-JAN-2000; 2000MO-US00219.
PR 06-JAN-2000; 2000MO-US00277.
PR 06-JAN-2000; 2000MO-US00376.
PR 11-FEB-2000; 2000MO-US0365.
PR 18-FEB-2000; 2000MO-US04341.
PR 22-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
XX
PA (GERTH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W,
PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
PI Stewart TR, Tunas D, Watanabe CK, Wood WI, Yan W;
XX
DR WPI; 2000-572271/53.
DR N-PSDB; AAC58626.
XX
PT
PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
XX
PS
XX Claim 33; Fig 96; 309p; English.

	CC	immunological diseases of the lung, and transplantation associated
	CC	diseases including graft rejection and graft-versus-host-disease.
	CC	AAC58397 to AAC58378 represent PCR primers and hybridisation probes used
	CC	in the isolation of human PRO sequences. AAC58579 to AAC58642 and
	CC	AAB33414 to AAB33477 represent human PRO polynucleotide and protein
	CC	sequences given in the exemplification of the present invention.
XX		
SQ	Sequence	243 AA;
	Query Match	100.0%; Score 1325; DB 21; Length 243;
	Best Local Similarity	100.0%; Pred. No. 5.3e-103;
	Matches 243; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY		
Dd	1	MRPLVLLVLGLAAGSPPLDNDKI ^{PSLCRGHPGAPGTGGHHSOQLPGRDGDGRDGARG} 60
OY	1	MRPLVLLVLGLAAGSPPLDNDKIPSLCPGHPLGFTGHHGSQGLPGRDGDGRDGARG
'Db	61	AFGEKGEGRGRLGPGRDPGPRBAGPAGPTGPRGECSSVPBRSAFSAKRSSRYPPSD
OY	61	AFGEKGEGRGRLGPGRDPGPRBAGPAGPTGPRGECSSVPBRSAFSAKRSSRYPPSD
Dd	121	APLPEDRVLVNVEOGHYDAVTGKTFCQVPGVYYFAVAHTVYRASLDFLYKNCESTASFPQ
OY	121	APLPEDRVLVNVEOGHYDAVTGKTFCQVPGVYYFAVAHTVYRASLDFLYKNCESTASFPQ
Dd	181	PFGGMPPRPAISLSGAMVRLBPEDDVWVOVGVDIYGIVASIKTDSFGSFLYSMDHSSP
OY	181	PFGGMPPRPAISLSGAMVRLBPEDDVWVOVGVDIYGIVASIKTDSFGSFLYSMDHSSP
Dd	241	VFA 243
OY	241	VFA 243
Dd	241	VFA 243
	RESULT 5	
	AAV71468	
	ID	AAV71468 standard; Protein; 243 AA.
XX		
AC	AAV71468;	
XX		
DT	08-NOV-2000	(first entry)
DB		
XX		
XX	Human PRO344 protein.	
KM	PRO344; DNA40592-1242; human; ATCC No: 209492; antiproliferative;	
KM	neoplastic cell growth inhibitor; cytostatic; treatment; cancer; tumour;	
KM	breast; prostate; colon; lung; renal; ovarian; central nervous system;	
KM	CNS; leukemia; melanoma; Expressed Sequence Tag; EST; secreted protein;	
KX	extracellular domain; ECD.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..15
FT	Modified-site	/label= Signal_peptide
FT	Protein	/note= "N-myristoylation site"
FT	Modified-site	/label= Mature_PRO344_protein
FT	Binding-site	/note= "N-myristoylation site"
FT	Modified-site	/note= "Cell attachment sequence"
FT	Modified-site	/note= "N-myristoylation site"
XX		
PN	MO200032778-A2.	
PD	08-JUN-2000.	
XX		
PF	30-NOV-1999;	99WO-US28409.
RR	01-DEC-1998;	98WO-US25108.

PR 16-DEC-1998; 98US-0112850.
 PR 22-DEC-1998; 98US-0113296.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 XX
 PA (GETH) GENENTECH INC.
 PI Chen J, Goddard A, Gurney AL, Hillan K, Napier W, Wood WI;
 XX WPI; 2000-412325/35.
 DR N-PDB; AAD01241.
 XX
 PT New composition useful for inhibiting neoplastic cell growth and for
 PT treating cancers, comprises PRO655, PRO344 or PRO364 polypeptide or
 PT their antagonists
 XX
 PS Claim 31; Fig 6; 108bp; English.
 CC
 CC The present sequence is the human PRO344 protein, encoded by the cDNA
 CC clone, designated as DNA40592-1242. It is isolated from human foetal
 CC lung tissue, cDNA library, identified using probes based on a consensus
 CC sequence DNA4398, derived from secreted protein extracellular domain
 CC (ECD) expressed sequence tag (EST). This clone is assigned ATCC deposit
 CC No. 209492. PRO344 functions as a neoplastic cell growth inhibitor and
 CC is used for treating tumours, using an effective amount of PRO655, PRO364
 CC and PRO344. This composition is especially useful for treatment of human
 CC cancers such as breast, prostate, colon, lung, renal, ovarian and CNS,
 CC leukemia and melanoma.
 CC
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 1325; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLGLAGSPPLDNDKIPSLCPGHPGLPGTGHGSGGLRGRDGRDAPG 60
 DB 1 MRPLVLLGLAGSPPLDNDKIPSLCPGHPGLPGTGHGSGGLRGRDGRDAPG 60
 QY 61 APGKGGGRGRLPGPRGDPGPRGEGAGPAGTGCAGGSCVPPRPAFAKSESSESVPPSD 120
 DB 61 APGKGGGRGRLPGPRGDPGPRGEGAGPAGTGCAGGSCVPPRPAFAKSESSESVPPSD 120
 QY 121 APLPEDVVLNEQGHYDAVTKFTQVPGVYFAVHAATVVRASIQFDLVKNGESIASFPQ 180
 DB 121 APLPEDVVLNEQGHYDAVTKFTQVPGVYFAVHAATVVRASIQFDLVKNGESIASFPQ 180
 QY 181 PFGGMPKPAISLGGAMWRLPEPDQVWVQVGVGDIYIYASIKTSTSGFLVYSDMHSSP 240
 DB 181 PFGGMPKPAISLGGAMWRLPEPDQVWVQVGVGDIYIYASIKTSTSGFLVYSDMHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 6
 AAY93688
 ID AAY93688 standard; Protein; 243 AA.
 AC AAY93688;
 XX
 DT 03-OCT-2000 (first entry)
 XX
 DE Amino acid sequence of novel polypeptide PRO344.
 XX
 KM PRO201; PRO22; PRO227; PRO1265; PRO344; PRO343; PRO347; PRO357;
 KM PRO715; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell;
 KM tumorigenesis; cancer; neoplastic cell growth; cell proliferation.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT Peptide 1..15
 FT /note= "signal peptide"
 FT Modified-site 11..17
 FT /note= "N-myristoylation site"
 FT Modified-site 68..74
 FT /note= "N-myristoylation site"
 FT Region 77..80
 FT /note= "cell attachment sequence"
 FT Modified-site 216..222
 FT /note= "N-myristoylation site"
 XX
 PN W0200037640-A2.
 XX
 XX 29-JUN-2000.
 PD
 XX
 PF 16-DEC-1999; 99WO-US30095.
 XX
 PR 22-DEC-1998; 98US-0113296.
 PR 08-MAR-1999; 99WO-US05028.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28565.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;
 PI Wood WI;
 XX WPI; 2000-452188/39.
 DR N-PDB; AAA46907.
 XX
 PT New anti-polypeptide antibody useful in the treatment and diagnosis of
 PT neoplastic cell growth and proliferation -
 XX
 PS Claim 61; Fig 10; 220bp; English.
 CC
 CC The present sequence represents a novel human polypeptide. The
 CC specification describes novel polypeptides designated PRO201, PRO292,
 CC PRO327, PRO1265, PRO344, PRO343, PRO347, PRO715, PRO1017,
 CC PRO1112, PRO509, PRO853, and PRO882. These genes are amplified in
 CC the genome of tumour cells. The polypeptides are believed to contribute
 CC to tumorigenesis. The polypeptides are useful target for the
 CC identification of certain cancers, and may act as predictors of the
 CC prognosis of tumour treatment. Antibodies against these polypeptides
 CC are useful in the treatment and diagnosis of neoplastic cell growth
 CC and proliferation in mammals.
 CC
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 1325; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLGLAGSPPLDNDKIPSLCPGHPGLPGTGHGSGGLRGRDGRDAPG 60
 DB 1 MRPLVLLGLAGSPPLDNDKIPSLCPGHPGLPGTGHGSGGLRGRDGRDAPG 60
 QY 61 APGKGGGRGRLPGPRGDPGPRGEGAGPAGTGCAGGSCVPPRPAFAKSESSESVPPSD 120
 DB 61 APGKGGGRGRLPGPRGDPGPRGEGAGPAGTGCAGGSCVPPRPAFAKSESSESVPPSD 120
 QY 121 APLPEDVVLNEQGHYDAVTKFTQVPGVYFAVHAATVVRASIQFDLVKNGESIASFPQ 180
 DB 121 APLPEDVVLNEQGHYDAVTKFTQVPGVYFAVHAATVVRASIQFDLVKNGESIASFPQ 180
 QY 181 PFGGMPKPAISLGGAMWRLPEPDQVWVQVGVGDIYIYASIKTSTSGFLVYSDMHSSP 240
 DB 181 PFGGMPKPAISLGGAMWRLPEPDQVWVQVGVGDIYIYASIKTSTSGFLVYSDMHSSP 240

Oy 241 VFA 243
 Db 241 VFA 243

RESULT 7
 AAB01318
 ID AAB01318 standard; Protein; 243 AA.

XX AAB01318;
 AC 25-SEP-2000 (first entry)
 DT
 XX
 DE Human PRO344 polypeptide.

XX PRO; membrane bound protein; secreted protein; PRO357; PRO327;
 KW PRO243; PRO715; PRO241; PRO239; PRO233; PRO344; PRO347;
 KW PRO355; PRO353; PRO361; PRO365; transmembrane polypeptide;
 KW antibody; screening; detection; inhibition; probe; primer; human.

XX Homo sapiens.

XX Key
 FH Peptide
 FT Location/Qualifiers
 FT 1..15
 FT /label= Signal peptide
 FT Modified-site
 FT 11..17
 FT /note= "N-myristoylation site"
 FT Modified-site
 FT 68..74
 FT /note= "N-myristoylation site"
 FT Region
 FT 77..80
 FT /label= Cell attachment sequence
 FT Modified-site
 FT 216..222
 FT /note= "N-myristoylation site"

XX MO200032776-A2.
 XX 08-JUN-2000.
 XX 01-DEC-1999; 99WO-US28301.
 XX 01-DEC-1998; 98WO-US25108.
 XX 16-DEC-1998; 98US-0112850.
 XX 22-DEC-1998; 98US-0113296.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Botstein D, Eaton DL, Ferrara N, Filvaroff E,
 PI Gerritsen MB, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL,
 PI Hillan KJ, Kljavin IJ, Napier MA, Roy MA, Tumas D, Wood WT;
 XX WPI; 2000-413324/35.
 XX DR N-PSDB; AAA49560.
 XX New human nucleic acids encoding secreted and transmembrane
 PT polypeptides, designated as PRO polypeptides, useful as pharmaceutical
 PT and diagnostic agents
 XX Claim 12; Fig 18; 187pp; English.

XX New human nucleic acids encoding secreted and transmembrane
 CC polypeptides which are designated as PRO polypeptides are described
 CC The membrane-bound proteins have various industrial applications,
 CC including as pharmaceutical and diagnostic agents. The membrane-bound
 CC proteins can also be employed for screening of potential peptide or
 CC small molecule inhibitors of the relevant receptor/ligand interaction.
 CC Anti-PRO antibodies are useful for the affinity purification of PRO
 CC from recombinant cell culture or natural sources.

XX Sequence 243 AA;

Query Match 100.0%; Score 1325; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHGILPTPGHHGSGQLPGRDGRDGDAG 60
 Db 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHGILPTPGHHGSGQLPGRDGRDGDAG 60

Oy 61 APGEKGGGRGRLPGPPGDDPGPRGAGPAGPTGAGGECVPPRASFSAKRESRPVPPSD 120
 Db 61 APGEKGGGRGRLPGPPGDDPGPRGAGPAGPTGAGGECVPPRASFSAKRESRPVPPSD 120

Oy 121 APLPFDRVLVNEQGHYAVTGKFTQVPGVYFPAVHAIVYRASIQFDLVKNGESTIASFFQ 180
 Db 121 APLPFDRVLVNEQGHYAVTGKFTQVPGVYFPAVHAIVYRASIQFDLVKNGESTIASFFQ 180

Oy 181 PFGGMPYRPASISGAMVRLBEPDQVWVQVGVGYIYIYASIKTUSTPSGFLVSDMHSSP 240
 Db 181 PFGGMPYRPASISGAMVRLBEPDQVWVQVGVGYIYIYASIKTUSTPSGFLVSDMHSSP 240

Oy 241 VFA 243
 Db 241 VFA 243

RESULT 8
 AAU12352
 ID AAU12352 standard; Protein; 243 AA.

XX AAU12352;
 AC 24-OCT-2001 (first entry)
 DT
 XX
 DE Human PRO344 polypeptide sequence.

XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor- α 1; TNF- α 1;
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIA; gene therapy.

XX Homo sapiens.
 XX MO200140466-A2.
 XX 07-JUN-2001.
 XX 01-DEC-2000; 2000WO-US32678.
 XX 01-DEC-1999; 99WO-US28301.
 XX 01-DEC-1999; 99WO-US28634.
 XX 02-DEC-1999; 99WO-US28551.
 XX 02-DEC-1999; 99WO-US28564.
 XX 02-DEC-1999; 99WO-US28565.
 XX 09-DEC-1999; 99US-0170262.
 XX 16-DEC-1999; 99WO-US30911.
 XX 20-DEC-1999; 99WO-US30999.
 XX 30-DEC-1999; 99WO-US31243.
 XX 06-JAN-2000; 2000WO-US00277.
 XX 06-JAN-2000; 2000WO-US00376.
 XX 11-FEB-2000; 2000WO-US03565.
 XX 18-FEB-2000; 2000WO-US04341.
 XX 18-FEB-2000; 2000WO-US04342.
 XX 22-FEB-2000; 2000WO-US04414.
 XX 24-FEB-2000; 2000WO-US04914.
 XX 24-FEB-2000; 2000WO-US05004.
 XX 01-MAR-2000; 2000WO-US05601.
 XX 20-MAR-2000; 2000WO-US07377.
 XX 21-MAR-2000; 2000WO-US07532.
 XX 30-MAR-2000; 2000WO-US08439.
 XX 17-MAY-2000; 2000WO-US13705.
 XX 22-MAY-2000; 2000WO-US14042.
 XX 30-MAY-2000; 2000WO-US14941.
 XX 02-JUN-2000; 2000WO-US15264.
 XX 10-NOV-2000; 2000WO-US30873.

PA (GETH) GENENTECH INC.
 XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gunney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI, 2001-408281/43.
 DR N-PSDB; AAS21424.
 XX
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 XX Claim 12; Fig 362; 813p; English.
 XX
 XX AAU2172-AAU21446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumor necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 CC
 XX Sequence 243 AA;
 SQ
 Query Match 100.0%; Score 1325; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLVLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPG 60
 DB 1 MRPLVLLVLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPG 60
 QY 61 APGEKGGGRGRLPGPRGDPGRGAGPAGTGPAGGCSVPPRSASFSAKRSSESVPPSPD 120
 DB 61 APGEKGGGRGRLPGPRGDPGRGAGPAGTGPAGGCSVPPRSASFSAKRSSESVPPSPD 120
 QY 121 APLEPDRVLVNEOGHYDAVTGKFTCOVGVYVFAVHAIVYRASLOPDLVKNGESIASFFQ 180
 DB 121 APLEPDRVLVNEOGHYDAVTGKFTCOVGVYVFAVHAIVYRASLOPDLVKNGESIASFFQ 180
 QY 181 FFGGMPKPRASISGAMWRLPEPDQVWVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
 DB 181 FFGGMPKPRASISGAMWRLPEPDQVWVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 9
 ID AAB65815
 XX AAB65815 standard; Protein; 243 AA.
 XX
 AC AAB65815;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DB Human TANGO 253 SEQ ID NO: 3.

XX
 KW Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
 KW INTERCEPT 256; coronary disorder; olfactory disorder;
 KW neurological disorder; pulmonary disorder; immunological disorder;
 KW developmental disorder; kidney disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200078808-A1.
 XX
 XX 28-DEC-2000.
 XX
 XX 19-JUN-2000; 2000WO-US16883.
 XX
 XX 18-JUN-1999; 99US-0336536.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 PI Leiby KR, McKay C, Bossone S;
 PI WPI, 2001-050109/06.
 DR
 XX New nucleic acids for treating diseases and disorders, e.g.
 PT atherosclerosis, infection, autoimmune diseases, obesity, ear
 PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
 PT sclerosis and asthma
 XX
 XX Claim 9; Page 211-212; 332p; English.
 XX
 XX The present invention provides the protein and coding sequences of the
 CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
 CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
 CC coronary, pulmonary, olfactory, immunological, neurological,
 CC developmental and kidney disorders.
 CC
 XX Sequence 243 AA;
 SQ
 Query Match 100.0%; Score 1325; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLVLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPG 60
 DB 1 MRPLVLLVLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPG 60
 QY 61 APGEKGGGRGRLPGPRGDPGRGAGPAGTGPAGGCSVPPRSASFSAKRSSESVPPSPD 120
 DB 61 APGEKGGGRGRLPGPRGDPGRGAGPAGTGPAGGCSVPPRSASFSAKRSSESVPPSPD 120
 QY 121 APLEPDRVLVNEOGHYDAVTGKFTCOVGVYVFAVHAIVYRASLOPDLVKNGESIASFFQ 180
 DB 121 APLEPDRVLVNEOGHYDAVTGKFTCOVGVYVFAVHAIVYRASLOPDLVKNGESIASFFQ 180
 QY 181 FFGGMPKPRASISGAMWRLPEPDQVWVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
 DB 181 FFGGMPKPRASISGAMWRLPEPDQVWVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 10
 ID AAB49593
 XX AAB49593 standard; Protein; 243 AA.
 XX
 AC AAB49593;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DB Human adipocyte complement related protein homolog zs1g39.
 XX
 KW Human; zacrps; gene therapy; complement inhibition; C1q domain;

KW adipocyte complement related protein homolog;
 KW inflammation; hormone secretion; inositol phosphate; arachidonate;
 KW phospholipase C activation; gastric emptying; neutrophil activation;
 KW superoxide anion production; antimicrobial; acute vascular injury;
 KW wound healing; zsig39.

OS Homo sapiens.

XX MO200073444-A1.

XX 07-DEC-2000.

XX 18-MAY-2000; 2000MO-US13608.

XX 27-MAY-1999; 99US-0321372.

XX (ZYMO) ZYMOGENETICS INC.

XX Piddington CS, Sheppard FO;

XX WPI; 2001-061531/07.

PT Novel adipocyte complement related protein homolog, ZACRP5, useful for
 PT diagnosing and treating inflammation, vascular injury microbial
 PT infections, and in wound healing

PS Disclosure; Fig 1; 121pp; English.

XX The present invention relates to human adipocyte complement related
 CC protein homolog, zacrp5 protein and coding sequence (see AAB49599 and
 CC AAC90045). ZACRP5 has a carboxyl-terminal C1q domain. The zacrp5 gene is
 CC located on human chromosome 16. zacrp5 gene and protein are useful for
 CC diagnosing and treating inflammations, for determining arterial
 CC remodeling, for modulating calcium ion concentration, hormone
 CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC zacrp5 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, for
 CC treating acute vascular injury, and for wound healing. The present
 CC sequence is human adipocyte complement related protein homolog, zsig39
 CC protein. This protein was used in a sequence homology alignment with
 CC zacrp5 protein.

XX Sequence 243 AA;

Query Match 100.0%; Score 1325; DB 22; Length 243;

Best Local Similarity 100.0%; Pred. No. 5.3e-103; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLVLLGLAAGSPPLDNNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGAAG 60

DB 1 MRPLVLLVLLGLAAGSPPLDNNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGAAG 60

QY 61 APGEKSGRPGPLPGRGDPGRGEAGPAGPTGAGCSVPKSAFSAKSESRRVPPSPD 120

DB 61 APGEKSGRPGPLPGRGDPGRGEAGPAGPTGAGCSVPKSAFSAKSESRRVPPSPD 120

QY 121 APLEPDRVLVNEQGHDAVTGKFTCOVPGVYFAVHAATVYRASLQFDLVNKGESIASFFQ 180

DB 121 APLEPDRVLVNEQGHDAVTGKFTCOVPGVYFAVHAATVYRASLQFDLVNKGESIASFFQ 180

QY 181 FFGWMPKPAASLGGAMVRLPEPDQVWQVGVGYIGIYASIKTDSFSGFLVSDMHSSP 240

DB 181 FFGWMPKPAASLGGAMVRLPEPDQVWQVGVGYIGIYASIKTDSFSGFLVSDMHSSP 240

QY 241 VFA 243

DB 241 VFA 243

RESULT 11

AAB49599

ID AAB49599 standard; Protein; 243 AA.

XX AAB49599;

AC 13-MAR-2001 (first entry)

DE Human adipocyte complement related protein homolog zsig39.

XX Human; zacrp6; gene therapy; complement inhibition; C1q domain;

KW adipocyte complement related protein homolog;

KW inflammation; hormone secretion; inositol phosphate; arachidonate;

KW phospholipase C activation; gastric emptying; neutrophil activation;

KW superoxide anion production; antimicrobial; acute vascular injury;

KW wound healing; zsig39.

OS Homo sapiens.

XX MO200073446-A2.

XX 07-DEC-2000.

XX 22-MAY-2000; 2000MO-US14024.

XX 27-MAY-1999; 99US-0321262.

XX (ZYMO) ZYMOGENETICS INC.

XX Piddington CS, Sheppard FO;

XX WPI; 2001-061532/07.

PT Novel adipocyte complement related protein homolog, ZACRP6, useful as
 PT modulators of neurotransmission and for treating disseminated
 PT intravascular coagulation, arteriosclerosis and acute vascular injury

PS Disclosure; Fig 1; 119pp; English.

XX The present invention relates to human adipocyte complement related
 CC protein homolog, ZACRP6 protein and coding sequence (see AAB49596 and
 CC AAC90051). ZACRP6 has a carboxyl-terminal C1q domain. The zacrp6 gene is
 CC located on human chromosome 21q. zacrp6 gene and protein are useful for
 CC diagnosing and treating inflammations, for determining arterial
 CC remodeling, for modulating calcium ion concentration, hormone
 CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC zacrp6 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, for
 CC treating acute vascular injury, disseminated intravascular coagulation,
 CC arteriosclerosis and for wound healing. The present sequence is human
 CC adipocyte complement related protein homolog zsig39. This protein was
 CC used in a sequence homology comparison with ZACRP6 protein.

XX Sequence 243 AA;

Query Match 100.0%; Score 1325; DB 22; Length 243;

Best Local Similarity 100.0%; Pred. No. 5.3e-103; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLVLLGLAAGSPPLDNNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGAAG 60

DB 1 MRPLVLLVLLGLAAGSPPLDNNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGAAG 60

QY 61 APGEKSGRPGPLPGRGDPGRGEAGPAGPTGAGCSVPKSAFSAKSESRRVPPSPD 120

DB 61 APGEKSGRPGPLPGRGDPGRGEAGPAGPTGAGCSVPKSAFSAKSESRRVPPSPD 120

QY 121 APLEPDRVLVNEQGHDAVTGKFTCOVPGVYFAVHAATVYRASLQFDLVNKGESIASFFQ 180

DB 121 APLEPDRVLVNEQGHDAVTGKFTCOVPGVYFAVHAATVYRASLQFDLVNKGESIASFFQ 180

QY 181 FFGWMPKPAASLGGAMVRLPEPDQVWQVGVGYIGIYASIKTDSFSGFLVSDMHSSP 240

Db 181 FFGMPKPSLSSGAMVRLPEPDQVWVQVGVGDIYIGIYASIKTDSITSGFLVYSMDHSSP 240
 Qy 241 VFA 243
 Db 241 VFA 243

RESULT 12

ID AAB65891 standard; Protein; 243 AA.

AC AAB65891;

DT 28-MAR-2001 (first entry)

DE Human secreted protein related protein SEQ ID NO: 108.

KM Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;

KW INTERCEPT 258; coronary disorder; olfactory disorder;

KM neurological disorder; pulmonary disorder; immunological disorder;

OS Homo sapiens.

PN WO200078808-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16883.

PR 18-JUN-1999; 99US-0336536.

PA (MILL-) MILLENNIUM PHARM INC.

PI Leiby KR, McKay C, Bossone S;

DR WPI; 2001-050109/06.

PT New nucleic acids for treating diseases and disorders, e.g.

PT atherosclerosis, infection, autoimmune diseases, obesity, ear

PT disorders, brain disorders, tumors, diabetes, arthritis, multiple

PT sclerosis and asthma

PS Disclosure; Page 274; 332pp; English.

CC The present invention provides the protein and coding sequences of the

CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,

CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of

CC coronary, pulmonary, olfactory, immunological, neurological,

CC developmental and kidney disorders.

CC Sequence 243 AA;

Qy Query Match 99.8%; Score 1323; DB 22; Length 243;

Best Local Similarity 99.6%; Pred. No. 7.8e-103;

Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

Qy 241 VFA 243
 Db 241 VFA 243

RESULT 13

ID AAB65888 standard; Protein; 243 AA.

AC AAB65888;

DT 28-MAR-2001 (first entry)

DE Human secreted protein related protein SEQ ID NO: 102.

KM Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;

KW INTERCEPT 258; coronary disorder; olfactory disorder;

KM neurological disorder; pulmonary disorder; immunological disorder;

OS Homo sapiens.

PN WO200078808-A1.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US16883.

PR 18-JUN-1999; 99US-0336536.

PA (MILL-) MILLENNIUM PHARM INC.

PI Leiby KR, McKay C, Bossone S;

DR WPI; 2001-050109/06.

PT New nucleic acids for treating diseases and disorders, e.g.

PT atherosclerosis, infection, autoimmune diseases, obesity, ear

PT disorders, brain disorders, tumors, diabetes, arthritis, multiple

PT sclerosis and asthma

PS Disclosure; Page 270-271; 332pp; English.

CC The present invention provides the protein and coding sequences of the

CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,

CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of

CC coronary, pulmonary, olfactory, immunological, neurological,

CC developmental and kidney disorders.

CC Sequence 243 AA;

Qy Query Match 99.7%; Score 1321; DB 22; Length 243;

Best Local Similarity 99.6%; Pred. No. 1.1e-102;

Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGRHHSOGLPGRDGRDGRDARG 60

Sat Jun 21 17:32:37 2003

us-09-944-403-42.rag

Page 11

Search completed: June 20, 2003, 11:23:01
job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:25:08 ; Search time 423 Seconds
(without alignments)
62.161 Million cell updates/sec

Title: US-09-944-403-42
Perfect score: 1325
Sequence: 1 MPRLVTLILGLAGSPPLD.....DSTFGSLVYSDWHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:*\n2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB pep:*\n3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:*\n4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*\n5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*\n6: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*\n7: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*\n8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB pep:*\n9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep:*\n10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:*\n11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:*\n12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:*\n13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:*\n14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1325	100.0	243	US-09-944-413-42	Sequence 42, App1
2	1325	100.0	243	US-09-944-403-42	Sequence 42, App1
3	1325	100.0	243	US-09-944-896-42	Sequence 42, App1
4	1325	100.0	243	US-09-944-944-42	Sequence 42, App1
5	1325	100.0	243	US-09-944-907-42	Sequence 42, App1
6	1325	100.0	243	US-09-944-929-42	Sequence 42, App1
7	1325	100.0	243	US-10-028-072-362	Sequence 362, App
8	1325	100.0	243	US-10-121-049-362	Sequence 362, App
9	1325	100.0	243	US-10-123-904-362	Sequence 362, App
10	1325	100.0	243	US-10-140-470-362	Sequence 362, App
11	1325	100.0	243	US-09-796-753-68	Sequence 68, App1
12	1325	100.0	243	US-10-175-746-362	Sequence 362, App
13	1325	100.0	243	US-10-176-918-362	Sequence 362, App
14	1325	100.0	243	US-10-176-921-362	Sequence 362, App
15	1325	100.0	243	US-10-137-865-362	Sequence 362, App
16	1325	100.0	243	US-10-140-474-362	Sequence 362, App
17	1325	100.0	243	US-10-142-431-362	Sequence 362, App
18	1325	100.0	243	US-10-143-114-362	Sequence 362, App
19	1325	100.0	243	US-10-140-002-362	Sequence 362, App

20 1325 100.0 243 9 US-10-142-419-362 Sequence 362, App
21 1325 100.0 243 9 US-10-123-262-362 Sequence 362, App
22 1325 100.0 243 9 US-10-142-423-362 Sequence 362, App
23 1325 100.0 243 9 US-10-121-050-362 Sequence 362, App
24 1325 100.0 243 9 US-10-141-755-362 Sequence 362, App
25 1325 100.0 243 9 US-10-143-032-362 Sequence 362, App
26 1325 100.0 243 9 US-10-123-108-362 Sequence 362, App
27 1325 100.0 243 9 US-10-123-236-362 Sequence 362, App
28 1325 100.0 243 9 US-10-123-261-362 Sequence 362, App
29 1325 100.0 243 9 US-10-140-821-362 Sequence 362, App
30 1325 100.0 243 9 US-10-140-928-362 Sequence 362, App
31 1325 100.0 243 9 US-10-121-045-362 Sequence 362, App
32 1325 100.0 243 9 US-10-123-292-362 Sequence 362, App
33 1325 100.0 243 9 US-10-123-903-362 Sequence 362, App
34 1325 100.0 243 9 US-10-124-819-362 Sequence 362, App
35 1325 100.0 243 9 US-10-124-822-362 Sequence 362, App
36 1325 100.0 243 9 US-10-140-925-362 Sequence 362, App
37 1325 100.0 243 9 US-10-160-498-362 Sequence 362, App
38 1325 100.0 243 9 US-09-944-884-42 Sequence 42, App1
39 1325 100.0 243 9 US-10-121-041-362 Sequence 362, App
40 1325 100.0 243 9 US-10-121-047-362 Sequence 362, App
41 1325 100.0 243 9 US-10-123-215-362 Sequence 362, App
42 1325 100.0 243 9 US-10-123-902-362 Sequence 362, App
43 1325 100.0 243 9 US-10-123-908-362 Sequence 362, App
44 1325 100.0 243 9 US-10-123-909-362 Sequence 362, App
45 1325 100.0 243 9 US-10-123-909-362 Sequence 362, App

ALIGNMENTS

RESULT 1
US-09-944-413-42
Sequence 42, Application US/09944413
Patent No. US20020156004A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Batton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OR INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944, 413
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866, 028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067, 411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069, 334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/06335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069, 696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 694
PRIOR FILING DATE: December 16, 1997

PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020156004A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020156004A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-413-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPHGLPTGPHHGSGLPGRDGRDCAAG 60
|||||

Db 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPHGLPTGPHHGSGLPGRDGRDCAAG 60
QY 61 APGEKGEGRPQLPGRPDGPRGEGAGPAGTGPAGGCSVPFSAKSESSESVPPSPD 120
Db 61 APGEKGEGRPQLPGRPDGPRGEGAGPAGTGPAGGCSVPFSAKSESSESVPPSPD 120
QY 121 APLEFDRVLVNEQGHYAVNGKFTCOVPGVYFPAVHAITYRASIQPLVYNGESIASFPQ 180
Db 121 APLEFDRVLVNEQGHYAVNGKFTCOVPGVYFPAVHAITYRASIQPLVYNGESIASFPQ 180
QY 181 FFGMPKXASISGAMVRLBEPDQVWVGVDYIGIYASIKTDSFGFLVYSDWSSP 240
Db 181 FFGMPKXASISGAMVRLBEPDQVWVGVDYIGIYASIKTDSFGFLVYSDWSSP 240
QY 241 VFA 243
Db 241 VFA 243

RESULT 2
US-09-944-403-42
Sequence 42, Application US/09944403
Patent No. US20020165143A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Batstein, David
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gottlsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OR INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,403
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998

PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020165143A1eember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020165143A1eember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-403-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLGTPGHHGSGOGLPGRDGRDGDGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLGTPGHHGSGOGLPGRDGRDGDGAPG 60
QY 61 APGEKGEGGRGLPGRDGRDGRGAGAGPTGPAEGESVPPKSAFSAKRSBSRVPPSD 120
DB 61 APGEKGEGGRGLPGRDGRDGRGAGAGPTGPAEGESVPPKSAFSAKRSBSRVPPSD 120
QY 121 APLPDRVLVNEOGHYDAVTKGTCQVGVYFAVHAIVYASLOPDLVKNGESIASFFQ 180
DB 121 APLPDRVLVNEOGHYDAVTKGTCQVGVYFAVHAIVYASLOPDLVKNGESIASFFQ 180
QY 181 FFGWPKPASLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVYSDMHSSP 240
DB 181 FFGWPKPASLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVYSDMHSSP 240

DB 181 FFGWPKPASLSCGAMVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 3
US-09-944-896-42
Sequence 42, Application US/09944896
Patent No. US20020168715A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gunney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Thomas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,896
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998

Page 4

APPLICANT: Eaton, Dan
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gerritsen, Mary
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul
 APPLICANT: Grimaldi, Christopher
 APPLICANT: Guirney, Auecin
 APPLICANT: Hillan, Kenneth
 APPLICANT: Kijavin, Ivar
 APPLICANT: Napier, Mary
 APPLICANT: Roy, Margaret
 APPLICANT: Tumas, Daniel
 APPLICANT: Wood, William
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P2548P1C1
 FILE OF INVENTION: ACIDS ENCODING THE SAME
 CURRENT APPLICATION NUMBER: US/09/944, 944
 CURRENT FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: 09/866, 028
 PRIOR FILING DATE: 2001-05-25
 PRIOR APPLICATION NUMBER: 60/067, 411
 PRIOR FILING DATE: December 3, 1997
 PRIOR APPLICATION NUMBER: 60/069, 334
 PRIOR FILING DATE: December 11, 1997
 PRIOR APPLICATION NUMBER: 60/069, 335
 PRIOR FILING DATE: December 11, 1997
 PRIOR APPLICATION NUMBER: 60/069, 278
 PRIOR FILING DATE: December 11, 1997
 PRIOR APPLICATION NUMBER: 60/069, 425
 PRIOR FILING DATE: December 12, 1997
 PRIOR APPLICATION NUMBER: 60/069, 696
 PRIOR FILING DATE: December 16, 1997
 PRIOR APPLICATION NUMBER: 60/069, 694
 PRIOR FILING DATE: December 16, 1997
 PRIOR APPLICATION NUMBER: 60/069, 702
 PRIOR FILING DATE: December 16, 1997
 PRIOR APPLICATION NUMBER: 60/069, 870
 PRIOR FILING DATE: December 17, 1997
 PRIOR APPLICATION NUMBER: 60/069, 873
 PRIOR FILING DATE: December 17, 1997

APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gertitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Nadier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,944
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,596
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,286
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020173463A1ember 30, 1999

PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020173463A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-944-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKTPSLCPGHPGLPCTGHHGSOGLPGRDGRDGDGARG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKTPSLCPGHPGLPCTGHHGSOGLPGRDGRDGDGARG 60
QY 61 APGEKGGGRPGLPGRDGPGRGAGPAGPTGPAEGCSVPPRSASFSAKRSRRVPPSD 120
DB 61 APGEKGGGRPGLPGRDGPGRGAGPAGPTGPAEGCSVPPRSASFSAKRSRRVPPSD 120
QY 121 APLPDRVLVNEQGHYDAVTGKFTCQVPGVYFAVHATVYRASLOFDLVKNGESTIASFPQ 180
DB 121 APLPDRVLVNEQGHYDAVTGKFTCQVPGVYFAVHATVYRASLOFDLVKNGESTIASFPQ 180
QY 181 FFGGMPKPAASLSGGAMVLEBEDQWVGVGDYIGIYASIKTSTFSGFLVYSWMHSP 240
DB 181 FFGGMPKPAASLSGGAMVLEBEDQWVGVGDYIGIYASIKTSTFSGFLVYSWMHSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 5

US-09-944-907-42
Sequence 42, Application US/09944907
Publication No. US20020198147A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret

APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACTS ENCODING THE SAME
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,907
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-907-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKTPSLCPGHPGLPCTGHHGSOGLPGRDGRDGDGARG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKTPSLCPGHPGLPCTGHHGSOGLPGRDGRDGDGARG 60
QY 61 APGEKGGGRPGLPGRDGPGRGAGPAGPTGPAEGCSVPPRSASFSAKRSRRVPPSD 120
DB 61 APGEKGGGRPGLPGRDGPGRGAGPAGPTGPAEGCSVPPRSASFSAKRSRRVPPSD 120
QY 121 APLPDRVLVNEQGHYDAVTGKFTCQVPGVYFAVHATVYRASLOFDLVKNGESTIASFPQ 180
DB 121 APLPDRVLVNEQGHYDAVTGKFTCQVPGVYFAVHATVYRASLOFDLVKNGESTIASFPQ 180
QY 181 FFGGMPKPAASLSGGAMVLEBEDQWVGVGDYIGIYASIKTSTFSGFLVYSWMHSP 240
DB 181 FFGGMPKPAASLSGGAMVLEBEDQWVGVGDYIGIYASIKTSTFSGFLVYSWMHSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 6

US-09-944-929-42
Sequence 42, Application US/09944929
Publication No. US20020197612A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Napier, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACTS ENCODING THE SAME
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,929
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT

ORGANISM: Homo Sapien
US-09-944-929-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,36-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLVLLGLAAGSPPLDNDKIPSLCPGHHGLFETGHHGSGQLPRDGDGDGDGAP 60
DB 1 MRPLVLLVLLGLAAGSPPLDNDKIPSLCPGHHGLFETGHHGSGQLPRDGDGDGDGAP 60
QY 61 APGEKGGRRPGLPGRGDPGRGAGAPGPGAGCSVPPRSAPSAKSSRRVPPSD 120
DB 61 APGEKGGRRPGLPGRGDPGRGAGAPGPGAGCSVPPRSAPSAKSSRRVPPSD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFPQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFPQ 180
QY 181 FFGGMPKPAISGGAMVRLPEPDQYVQVGVGDYIGIYASIKTDSITPSGFLVYSMDHSSP 240
DB 181 FFGGMPKPAISGGAMVRLPEPDQYVQVGVGDYIGIYASIKTDSITPSGFLVYSMDHSSP 240
QY 241 VPA 243
DB 241 VPA 243

RESULT 7

US-10-028-072-362
Sequence 362, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunes, Daniel
APPLICANT: Tunes, Daniel
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072220
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20

;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-02-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081695
;; PRIOR FILING DATE: 1998-04-14
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081818
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082999
;; PRIOR FILING DATE: 1998-04-24
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085149
;; PRIOR FILING DATE: 1998-05-12
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086414
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086430
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088730
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088741
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090538
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRPLVLLILGLAAGSPPLDDNKIPSLCRGHPGLPCTPGHSGQLPGRDGRDGPAG	60
Db	1	MRPLVLLILGLAAGSPPLDDNKIPSLCRGHPGLPCTPGHSGQLPGRDGRDGPAG	60
Qy	61	APGKGGGRPGLRGPRGDPGPRGEGAPGPTGPGAGCSVPSPSARSRSPVPPSPD	120
Db	61	APGKGGGRPGLRGPRGDPGPRGEGAPGPTGPGAGCSVPSPSARSRSPVPPSPD	120
Qy	121	APLPFRVLVNEQGHDAVTGKFTCOVGVYFAVHATVYRASLOPDLVNGESIASFFQ	180
Db	121	APLPFRVLVNEQGHDAVTGKFTCOVGVYFAVHATVYRASLOPDLVNGESIASFFQ	180
Qy	181	FFGMPKPSLSGQAVRLPEPDQVWVGVDYIGIVASIKTDSFGFLVYSWMHSP	240
Db	181	FFGMPKPSLSGQAVRLPEPDQVWVGVDYIGIVASIKTDSFGFLVYSWMHSP	240
Qy	241	VFA 243	
Db	241	VFA 243	

RESULT 8

US-10-121-049-362

; Sequence 362, Application US/10121049
; Publication No. US2003002239A1

GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Beresini, Maureen
;; APPLICANT: DeForge, Laura
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Sherwood, Steven
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3330RIC17
;; CURRENT APPLICATION NUMBER: US/10/121,049
;; CURRENT FILING DATE: 2002-04-12
;; PRIOR APPLICATION removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 362
;; LENGTH: 243
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-121-049-362

Query Match 100.0%; Score 1325; DB 9; Length 243;

Best Local Similarity 100.0%; Pred. No. 2.3e-89; Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGGLPGRDGRDGDGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGGLPGRDGRDGDGAPG 60
QY 61 APGEKGEGRPGLPGRGDPGRGEGAGPTGPAECVPPRSAFSAKRSRVPSPSD 120
DB 61 APGEKGEGRPGLPGRGDPGRGEGAGPTGPAECVPPRSAFSAKRSRVPSPSD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATVYRASLOPDLVKNESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATVYRASLOPDLVKNESIASFFQ 180
QY 181 FFGGMPKPASLSGAMVRLPEPDQVWVQGVGDYIGIVASIKTDSFSGFLVYSDWHSSP 240
DB 181 FFGGMPKPASLSGAMVRLPEPDQVWVQGVGDYIGIVASIKTDSFSGFLVYSDWHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 9

US-10-123-904-362
Sequence 362, Application US/10123904
Publication No. US20030022328A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C54
CURRENT APPLICATION NUMBER: US/10/123,904
PRIORITY FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-904-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGGLPGRDGRDGDGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGGLPGRDGRDGDGAPG 60
QY 61 APGEKGEGRPGLPGRGDPGRGEGAGPTGPAECVPPRSAFSAKRSRVPSPSD 120
DB 61 APGEKGEGRPGLPGRGDPGRGEGAGPTGPAECVPPRSAFSAKRSRVPSPSD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATVYRASLOPDLVKNESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATVYRASLOPDLVKNESIASFFQ 180

DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATVYRASLOPDLVKNESIASFFQ 180

QY 181 FFGGMPKPASLSGAMVRLPEPDQVWVQGVGDYIGIVASIKTDSFSGFLVYSDWHSSP 240
DB 181 FFGGMPKPASLSGAMVRLPEPDQVWVQGVGDYIGIVASIKTDSFSGFLVYSDWHSSP 240

QY 241 VFA 243
DB 241 VFA 243

RESULT 10

US-10-140-470-362
Sequence 362, Application US/10140470
Publication No. US20030022331A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C160
CURRENT APPLICATION NUMBER: US/10/140,470
PRIORITY FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-470-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGGLPGRDGRDGDGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTGHHGSGGLPGRDGRDGDGAPG 60
QY 61 APGEKGEGRPGLPGRGDPGRGEGAGPTGPAECVPPRSAFSAKRSRVPSPSD 120
DB 61 APGEKGEGRPGLPGRGDPGRGEGAGPTGPAECVPPRSAFSAKRSRVPSPSD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATVYRASLOPDLVKNESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATVYRASLOPDLVKNESIASFFQ 180
QY 181 FFGGMPKPASLSGAMVRLPEPDQVWVQGVGDYIGIVASIKTDSFSGFLVYSDWHSSP 240
DB 181 FFGGMPKPASLSGAMVRLPEPDQVWVQGVGDYIGIVASIKTDSFSGFLVYSDWHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 11
US-09-796-753-68

```

; Sequence 68, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 68
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-68

```

```

Query Match      100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 MRPLVLLILGLAAGSPPLDNNKIPSLCPGHGPGLPCTPGHHSQGLPGRDGRDGRDGPAG 60
|||||

```

```

Db      1 MRPLVLLILGLAAGSPPLDNNKIPSLCPGHGPGLPCTPGHHSQGLPGRDGRDGRDGPAG 60
Qy      61 APGEKGEGRPGLPGRDGPGRGAGPAGTGPAGCSVPSPASAKRSRSRVPPSD 120
Db      61 APGEKGEGRPGLPGRDGPGRGAGPAGTGPAGCSVPSPASAKRSRSRVPPSD 120
Qy      121 APLEPDRVLYNEQGHDAVTGKTCQVPYVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
Db      121 APLEPDRVLYNEQGHDAVTGKTCQVPYVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
Qy      181 FPGGMPKPASLSGAMVRLPEBDQVWVGVDYIGIYASIKTDSFTSGGLVYSDMHSSP 240
Db      181 FPGGMPKPASLSGAMVRLPEBDQVWVGVDYIGIYASIKTDSFTSGGLVYSDMHSSP 240
Qy      241 VFA 243
Db      241 VFA 243

```

RESULT 12

```

US-10-175-746-362
; Sequence 362, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330RIC353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-362

```

```

Query Match      100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MRPLVLLILGLAAGSPPLDNNKIPSLCPGHGPGLPCTPGHHSQGLPGRDGRDGRDGPAG 60
Db      1 MRPLVLLILGLAAGSPPLDNNKIPSLCPGHGPGLPCTPGHHSQGLPGRDGRDGRDGPAG 60
Qy      61 APGEKGEGRPGLPGRDGPGRGAGPAGTGPAGCSVPSPASAKRSRSRVPPSD 120
Db      61 APGEKGEGRPGLPGRDGPGRGAGPAGTGPAGCSVPSPASAKRSRSRVPPSD 120
Qy      121 APLEPDRVLYNEQGHDAVTGKTCQVPYVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
Db      121 APLEPDRVLYNEQGHDAVTGKTCQVPYVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
Qy      181 FPGGMPKPASLSGAMVRLPEBDQVWVGVDYIGIYASIKTDSFTSGGLVYSDMHSSP 240
Db      181 FPGGMPKPASLSGAMVRLPEBDQVWVGVDYIGIYASIKTDSFTSGGLVYSDMHSSP 240

```

QY 241 VFA 243
 |||
 Db 241 VFA 243

RESULT 13

US-10-176-918-362
 ; Sequence 362, Application US/10176918
 ; Publication No. US20030027275A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P330R1C382
 ; CURRENT FILING DATE: 2002-06-20
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 362
 ; LENGTH: 243
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-176-918-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
 Best Local Similarity 100.0%; Pred. No. 2.3e-89;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAPG 60
 |||
 Db 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAPG 60
 QY 61 APGEKGGRRPGLPGPRGDPGPRGAGPAGPTGAGCSVPPRSAFSAKSESSESVPPSPD 120
 |||
 Db 61 APGEKGGRRPGLPGPRGDPGPRGAGPAGPTGAGCSVPPRSAFSAKSESSESVPPSPD 120
 QY 121 APLPFDRLVNBQGHYDAVTGKFTCOVPGVYFPAVHATVYRASIQFDLVNKGESIASFPQ 180
 |||
 Db 121 APLPFDRLVNBQGHYDAVTGKFTCOVPGVYFPAVHATVYRASIQFDLVNKGESIASFPQ 180
 QY 181 FFGMPKPRASISGGMVRLBEPDQVWQVGVGYIGIYASIKTSTSGFLVYSDMHSSP 240
 |||
 Db 181 FFGMPKPRASISGGMVRLBEPDQVWQVGVGYIGIYASIKTSTSGFLVYSDMHSSP 240
 QY 241 VFA 243
 |||
 Db 241 VFA 243

RESULT 14

US-10-176-921-362
 ; Sequence 362, Application US/10176921
 ; Publication No. US20030027276A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P330R1C288
 ; CURRENT FILING DATE: 2002-06-20
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 362
 ; LENGTH: 243
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-176-921-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
 Best Local Similarity 100.0%; Pred. No. 2.3e-89;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAPG 60
 |||
 Db 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPGTPGHHGSGQLPGRDGRDGDAPG 60
 QY 61 APGEKGGRRPGLPGPRGDPGPRGAGPAGPTGAGCSVPPRSAFSAKSESSESVPPSPD 120
 |||
 Db 61 APGEKGGRRPGLPGPRGDPGPRGAGPAGPTGAGCSVPPRSAFSAKSESSESVPPSPD 120
 QY 121 APLPFDRLVNBQGHYDAVTGKFTCOVPGVYFPAVHATVYRASIQFDLVNKGESIASFPQ 180
 |||
 Db 121 APLPFDRLVNBQGHYDAVTGKFTCOVPGVYFPAVHATVYRASIQFDLVNKGESIASFPQ 180
 QY 181 FFGMPKPRASISGGMVRLBEPDQVWQVGVGYIGIYASIKTSTSGFLVYSDMHSSP 240
 |||
 Db 181 FFGMPKPRASISGGMVRLBEPDQVWQVGVGYIGIYASIKTSTSGFLVYSDMHSSP 240
 QY 241 VFA 243
 |||
 Db 241 VFA 243

RESULT 15

US-10-137-865-362
 ; Sequence 362, Application US/10137865
 ; Publication No. US20030032155A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C154
CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-865-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPIGTRGHHSQGLPRDGRDGRDGPAG 60
DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPIGTRGHHSQGLPRDGRDGRDGPAG 60
QY 61 APGEKGGGRPGLPGRGDPGPRGAGAPGPTGPGAGECSVPRSAFSAKRSRVPSPSD 120
DB 61 APGEKGGGRPGLPGRGDPGPRGAGAPGPTGPGAGECSVPRSAFSAKRSRVPSPSD 120
QY 121 APLPFDRLVNEQGHYDAVTGKPTQVGVYYPFAVHATVYRASLQFDLVNNGESIASFFQ 180
DB 121 APLPFDRLVNEQGHYDAVTGKPTQVGVYYPFAVHATVYRASLQFDLVNNGESIASFFQ 180
QY 181 FFGWMPKPSLISGAMVRLSPEDQVWVGVGDIYIGIYASIKTDSFGFLVYSDWHSSP 240
DB 181 FFGWMPKPSLISGAMVRLSPEDQVWVGVGDIYIGIYASIKTDSFGFLVYSDWHSSP 240
QY 241 VFA 243
DB 241 VFA 243

Search completed: June 20, 2003, 11:39:57
Job time : 424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 11:18:01 ; Search time 44 Seconds

(without alignments)
530.924 Million cell updates/sec

Title: US-09-944-403-42

Perfect score: 1325

Sequence: 1 NRPPLVLLILGLAASPPLD.....DSTFGFLVYSDMHSVPFA 243

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1207	91.1	219	2 T14782	hypothetical prote
2	424	32.0	244	2 JC4708	gelatin-binding 28
3	418.5	31.6	680	1 CGH01D	collagen alpha 1(X
4	416.5	31.4	674	2 S23297	collagen alpha 1(X
5	412	31.1	246	2 S29328	complement subcomp
6	411.5	31.1	674	2 S13301	collagen alpha 1(X
7	408	30.8	680	2 S31216	collagen alpha 1(X
8	399	30.1	245	1 C1H00C	complement subcomp
9	374	28.2	744	1 A34246	collagen alpha 1(V
10	370	27.9	744	1 S23298	collagen alpha 1(V
11	363.5	27.4	743	1 S23779	collagen alpha 1(V
12	363	27.4	635	1 A57131	collagen alpha 2(V
13	362	27.3	253	1 C1H00B	complement subcomp
14	360.5	27.2	253	2 S15435	complement protein
15	359	27.1	744	2 S15435	collagen alpha 1(V
16	351	26.5	253	2 A55797	complement C1q B c
17	316	23.8	423	2 A55797	collagen precursor
18	314	23.7	245	1 C1H00A	complement subcomp
19	314	23.7	245	2 S19018	complement subcomp
20	280.5	21.2	215	2 B48150	hibernation-relate
21	277.5	20.9	196	2 C48150	hibernation-relate
22	261.5	19.7	196	2 A48150	hibernation-relate
23	238.5	18.0	992	2 T08772	hypothetical prote
24	230.5	17.4	1049	1 CGB07S	collagen alpha 1(I
25	227.5	17.2	248	2 I51921	pulmonary surfacta
26	225.5	17.0	248	1 LNH0P1	pulmonary surfacta
27	225	17.0	325	2 T32248	hypothetical prote
28	224	16.9	636	2 S41067	collagen alpha 1(I
29	223	16.8	1464	2 S59856	collagen alpha 1(I

30	222.5	16.8	1758	2 T29350	hypothetical prote
31	222.5	16.8	1759	2 T29351	collagen alpha 2(I
32	222	16.8	360	2 T37285	collagen dy-2 - C
33	221.5	16.7	170	2 B57131	collagen alpha 2(V
34	221	16.7	886	2 I50694	collagen alpha 1(I
35	221	16.7	1019	1 A32856	collagen alpha 1(V
36	220	16.6	248	1 LND0FS	pulmonary surfacta
37	220	16.6	341	2 T16296	hypothetical prote
38	220	16.6	380	2 T28888	cuticle collagen d
39	219.5	16.6	247	1 LNRBPS	pulmonary surfacta
40	219.5	16.6	381	2 T27806	hypothetical prote
41	219.5	16.6	673	1 CGB06C	collagen alpha 1(I
42	219.5	16.6	1466	1 CGH07L	collagen alpha 1(I
43	219	16.5	671	1 CGRT1S	collagen alpha 1(I
44	218.5	16.5	1453	2 S21626	collagen alpha 1(I
45	217	16.4	283	2 T29980	hypothetical prote

ALIGNMENTS

RESULT 1

T14782
hypothetical protein DKFZp586B0621.1 - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000

C/Accession: T14782

R/Ottewillder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999

A/Reference number: Z18184

A/Accession: T14782

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-219 <OTT>

A/Cross-references: EMBL:AL110261

A/Experimental source: adult uterus; clone DKFZp586B0621

C/Genetics:

A/Note: DKFZp586B0621.1

C/Superfamily: complement C1q carboxyl-terminal homology

Query Match	Score 1207; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.3e-77;	
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	25 PSICPGHPLGPTGHHGSGQLPGRDGRDGAFCAPGKSGRGLPGRGDPGRG 84
Db	1 PSICPGHPLGPTGHHGSGQLPGRDGRDGAFCAPGKSGRGLPGRGDPGRG 60
QY	85 EAGPAGPTGAGGCVPPPSAFSAKRSERVPSPDAPLPFDRLVNEGSHDVAVYKGT 120
Db	61 EAGPAGPTGAGGCVPPPSAFSAKRSERVPSPDAPLPFDRLVNEGSHDVAVYKGT 120
QY	145 QVPPGTYPRVAVNTVYRASLOPDLVNGKSIASFPGFGMPKPPASISGAMVRLPEDO 204
Db	121 QVPPGTYPRVAVNTVYRASLOPDLVNGKSIASFPGFGMPKPPASISGAMVRLPEDO 180
QY	205 VMVVGVDYIGIYASIKTDSFGFLVYSDMHSVPFA 243
Db	181 VMVVGVDYIGIYASIKTDSFGFLVYSDMHSVPFA 219

RESULT 2

JC4708
gelatin-binding 28K protein precursor - human

N/Alternate names: adipose specific collagen-like factor

C/Species: Homo sapiens (man)

C/Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Sep-1999

C/Accession: JC4708; JC4944

R/Maeda, K.; Okubo, K.; Shimomura, I.; Funahashi, T.; Matsuzawa, Y.; Matsubara, K.
Biochem. Biophys. Res. Commun. 221, 286-289, 1996

A/Title: cDNA cloning and expression of a novel adipose specific collagen-like factor

A/Reference number: JC4708; MUID:96224171; PMID:8619847

A/Accession: JC4708

A: Molecule type: mRNA
 A: Residues: 1-244 <MAE>
 A: Cross-references: DDBJ:D45371; NID:9871886; PIDN:BA008227.1; PID:9871887
 A: Experimental source: adipose tissue
 R: Nakano, Y.; Tobé, T.; Choi-Miyura, N.H.; Mazda, T.; Tomita, M.
 J. Biochem. 120, 803-812, 1996
 A: Title: Isolation and characterization of GBP28, a novel gelatin-binding protein purified from sheep placenta
 A: Reference number: J04944; MUID:97103474; PMID:8947845
 A: Accession: J04944
 A: Molecule type: protein
 A: Residues: 19-38; 93-100; 101-112; 135-149; 173-178 <NMX>
 A: Comment: This protein is an endogenous factor that binds with a collagen-like domain.
 C: Genetics:
 A: Gene: apm1
 C: Superfamily: unassigned collagens; complement C1q carboxyl-terminal homology
 C: Keywords: adipose tissue; glycoprotein; hydroxyproline
 F: 1-18/Domain: signal sequence #status predicted <SIG>
 F: 19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>
 F: 42-107/Region: collagen-like
 F: 114-241/Domain: complement C1q carboxyl-terminal homology <C1Q>
 F: 95/Modified site: 4-hydroxyproline (Pro) #status experimental
 F: 230/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.0%; Score 424; DB 2; Length 244;
 Best Local Similarity 40.5%; Pred. No. 7.7e-23;
 Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;

OY 6 VLLLLGLAGSPPLDNNKIPSLCPH-----PG-----LPSTPGHNSQ 44
 DB 7 VLLHLL-----FCHDETTTQSGVLLPLPKAGCTCMAMGIRGPHN 50

OY 45 GLPGRDGDGDGAPAFGEKGRGRLPGRPD-----PGRPRGAPGAPGPA 95
 DB 51 GAPGRDGD-----GTPEKGEKGDPLIGPKGIGETGVPAEGPFGIQRKGP 104

OY 96 GECVPRPSARSAKRSRVRPPSDAPLPDRVLNBECHDAVAGKTCQVPGVYFAY 155
 DB 105 GEGAVVVSAPSV-GLETVYVIP-NMPLRFKIFRYNQQNHYDGSFGKHCINPGLYFAY 162

OY 156 HATVRAALQFDVLVNGERS-IASFPQFGWPKPASLSEGMVRLPEPDQVVOV-GVGD 213
 DB 163 HITVYMKVAVSLPFXKDKAMLFETYQYQENNVDA--SGSVLLHLEVDQWLVQYGE 220

OY 214 YIGYASIKTGSTSGFLVYSD 235
 DB 221 RRGVYADNDNDSTFGFLYHD 242

RESULT 3
 CGHUID
 collagen alpha 1(X) chain precursor - human
 N: Alternate names: procollagen alpha 1(X) chain
 C: Species: Homo sapiens (man)
 C: Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text change 22-Jun-1999
 R: Accession: S26396; S30086; S158249; A43901; I51870; S21556
 R: Reichenberger, B.; Beier, F.; Lwaile, P.; Olsen, B.R.; von der Mark, K.; Bertling, W.
 FEBS Lett. 311, 305-310, 1992
 A: Title: Genomic organization and full-length cDNA sequence of human collagen X.
 A: Reference number: S26396; MUID:93012005; PMID:1397333
 A: Accession: S26396
 A: Molecule type: DNA
 A: Residues: 1-680 <RET>
 A: Cross-references: EMBL:X68952; EMBL:X72578; EMBL:X72579; EMBL:X72580; GB:S47714; GB:S47715; S. submitted to the EMBL Data Library, March 1992
 A: Reference number: S30085
 A: Accession: S30085
 A: Molecule type: DNA
 A: Residues: 1-52-680 <AP2>
 A: Cross-references: EMBL:X65120; NID:923129
 A: Note: The initial difference is probably due to translation of an intronic sequence
 R: Apte, S.; Mattei, M.G.; Olsen, B.R.
 FEBS Lett. 282, 393-396, 1991

A: Title: Cloning of human alpha-1(X) collagen DNA and localization of the COL10A1 gene
 A: Reference number: S15826; MUID:91243838; PMID:2037056
 A: Accession: S15826
 A: Molecule type: DNA
 A: Residues: 561-647, 'G', 649-666 <AP2>
 R: Thomas, J.T.; Cresswell, C.J.; Raab, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant
 Biochem. J. 280, 617-623, 1991
 A: Title: The human collagen X gene. Complete primary translated sequence and chromosomal location
 A: Reference number: S18249; MUID:92109659; PMID:1764025
 A: Accession: S18249
 A: Molecule type: DNA
 A: Residues: 1-26, 'T', 28-680 <THO>
 A: Cross-references: EMBL:X60382; NID:930094; PIDN:CAA4293.1; PID:930095
 A: Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 532-
 R: Reichenberger, B.; Aigner, T.; von der Mark, K.; Stoss, H.; Bertling, W.
 Dev. Biol. 148, 562-572, 1991
 A: Title: In situ hybridization studies on the expression of type X collagen in fetal
 A: Reference number: A43901; MUID:92077285; PMID:1743401
 A: Accession: A43901
 A: Molecule type: mRNA
 A: Residues: 547-656 <RB2>
 A: Cross-references: GB:M74050; GB:D57494; NID:9339884; PIDN:AAA61221.1; PID:9553796
 A: Note: sequence extracted from NCB1 backbone (NCBI:69012, NCBIP:69014)
 R: Wallis, G.A.; Raab, B.; Sweetman, W.A.; Thomas, J.T.; Super, M.; Evans, G.; Grant,
 Am. J. Hum. Genet. 54, 169-178, 1994
 A: Title: Amino acid substitutions of conserved residues in the carboxyl-terminal domain of Schmid.
 A: Reference number: I51870; MUID:94136476; PMID:8304336
 A: Accession: I51870
 A: Status: translated from GB/EMBL/DBJ
 A: Molecule type: mRNA
 A: Residues: 520-597, 'D', 599-680 <MAL>
 A: Cross-references: GB:S68331; NID:9545180; PIDN:AAC0615.1; PID:9545181
 A: Note: mutant sequence from patient with metaphyseal chondrodysplasia type Schmid
 C: Comment: a second mutant sequence with 614-Pro is also described
 C: and subsequently O-glycosylated.
 C: Genetics:
 A: Gene: GDB:COL10A1
 A: Cross-references: GDB:128635; OMIM:120110
 A: Map position: 6q21-6q22
 A: Intron: 52/1
 A: Note: a defect in this gene may cause Schmid metaphyseal chondrodysplasia
 C: Complex: type X collagen may be a homotrimer
 C: Function:
 A: Description: structural component of extracellular fibrous polymer specifically and
 be important for skeletogenesis
 C: Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
 C: Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyls
 F: 1-18/Domain: signal sequence #status predicted <SIG>
 F: 19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
 F: 19-56/Domain: amino-terminal nonhelical #status predicted <NC2>
 F: 57-519/Region: interrupted helical
 F: 520-680/Domain: amino-terminal nonhelical #status predicted <NC1>
 F: 553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>
 F: 617/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.6%; Score 418.5; DB 1; Length 680;
 Best Local Similarity 34.8%; Pred. No. 5.4e-22;
 Matches 98; Conservative 34; Mismatches 82; Indels 67; Gaps 6;

OY 15 GSPRLDNNKIPSLCPHGLPGRPHNGSGGLGDRDGRDAPGAPGKSGGGRGLP 74
 DB 401 GNPRLPGRKDPGVGGPGLPGRVGRAGAGMPGHNGEACPRGAPGIGTRGPIGPGRIP 460

OY 75 -----GPRDPPGRGA-----GPAPTGPAGCS 99
 DB 461 GPRSKDPPGPPRPGATATKGLNGPTPPPPGRRGSGRGLGPPGPPGQAV 520

OY 100 VP-----PRSAFSAKSSRRVPPSDAPLPDRVLN 132
 DB 521 MPGRFIRAGORPSLSTGPLVANSAGVTVGMSAFVILSKAY--PAIGTIPFDKILYNR 578

QY 15 GSPPLDNKIPSLCPGRHGLPGTRHHGSQGLPGRDGRDGDGA PGAGKEGEGRGLP 74
 DB 401 GNPGLPGRPKDPIAGSPGLPGVPGAGAKGVPGHNGEAGRGVGLPGTRGPIGPGRIP 460
 QY 75 -----GPRGDPGRGEA-----GPAGPTPGAGCS 99
 DB 461 GPRGSKDVGTPGPRGAGIAVKGLNGPTGPRGPRGRNGRNGRGLRPPRPPPGVYA 520
 QY 100 VP-----PRSAFSAKRSESRVPPSPDAPLPEDRVLVNQGHYDA 138
 DB 521 LPEDFVAGRGPRFVSANQGVGMVSAFTVILSKAY--PAIGTPIPEDKILYNKQHYDP 578
 QY 139 VTGFTQGVGVVYFAVHATVYRASLOFDLVKNGESIA-SFQPRFGMKPKASISGGMV 197
 DB 579 RTGFTCKIPGITYFYFHHKTHAWGLKNGTPVWYTYDEVYIKGLDQA--SGSAVI 636
 QY 198 LPEPDQVWVQGVGDYIGIYASIKTSTFGSFLV 232
 DB 637 DLTNDQVWVWQLPNAAGSNGLYSPRYVHSSFGSFLV 671

RESULT 7
 S31216
 collagen alpha 1(X) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
 C:Accession: S31216; S28807; S22215; S30127; 148299; S26397; S31830
 R:Kong, R.Y.C.; Khan, K.M.; Lau, R.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.; C
 Eur. J. Biochem. 213, 99-111, 1993
 A:Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp
 A:Reference number: S31216; MUID:93238750; PMID:8477738
 A:Accession: S31216
 A:Molecule type: DNA
 A:Residues: 1-680 <KON>
 A:Cross-references: EMBL:X21610; NID:949793; PIDN:CAA79736.1; PID:949794
 R:Elima, K.; Berola, I.; Rosati, R.; Meteseranta, M.; Garofalo, S.; de Cro
 Biochem. J. 289, 247-253, 1993
 A:Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp
 A:Reference number: S28807; MUID:93143676; PMID:8424763
 A:Accession: S28807
 A:Molecule type: DNA
 A:Residues: 1-285, 'A', 287-680 <ELI>
 A:Cross-references: EMBL:X67348; NID:950480; PIDN:CAA47763.1; PID:950481
 R:Elima, K.; Meteseranta, M.; Kallio, J.; Petäelä, M.; Berola, I.; Garofalo, S.; de Cro
 Biochim. Biophys. Acta 1130, 78-80, 1992
 A:Title: Specific hybridization probes for mouse alpha-2(IX) and alpha-1(X) collagen mRN
 A:Reference number: S22215; MUID:92182017; PMID:1543751
 A:Accession: S22215
 A:Molecule type: DNA
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-12, 'F', 14-26, 'S', 28-247, 'L', 249-285, 'A', 287-305, 'F', 307-416, 'S', 418-499, 'I
 R:Appe, S.S.; Olsen, B.R.
 Eur. J. Biochem. 206, 217-224, 1992
 A:Title: Cloning of the human and mouse type X collagen genes and mapping of the mouse t
 A:Reference number: 148299; MUID:92267014; PMID:1587721
 A:Accession: 148299
 A:Status: preliminary
 A:Molecule type: translated from GB/EMBL/DBJ
 A:Residues: 52-247, 'L', 249-285, 'A', 287-305, 'F', 307-416, 'S', 418-499, 'I', 501-566, 'C', 568, 'J
 A:Cross-references: EMBL:X65121; NID:950482; PIDN:CAA46237.1; PID:9667031
 R:Summers, T.A.; Irwin, M.H.; Mayne, R.; Ballan, G.
 J. Biol. Chem. 263, 581-587, 1988
 A:Title: Monoclonal antibodies to type X collagen. Biosynthetic studies using an antibody

A:Reference number: S26397; MUID:88087150; PMID:2826450
 A:Accession: S26397
 A:Molecule type: protein
 A:Residues: 'SDGYSQ', 24-26, 'KO' <SUM>
 C:Genetics:
 A:Gene: Coll10a-1
 A:Map position: 10
 A:Introns: 51/3
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolo
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
 F:553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 30.8%; Score 408; DB 2; Length 680;
 Best Local Similarity 33.4%; Pred. No. 2.9e-21;
 Matches 99; Conservative 34; Mismatches 75; Indels 88; Gaps 9;

QY 15 GSPPLDNKIPSLCPGRHGLP-----GPRGHH-----GSQ 44
 DB 392 GEPGLNGPK-----GNPGLGQKGDPEVGTGTPGLRGPVGVAKGVPGHNGEAGRGPR 445
 QY 45 GLPRGRDGRDGA PGAGKEGEGRGLPGR-----GPRGGEA----- 86
 DB 446 GIPGTRGPTGPRGVPGFPGSKGDPNGCAPGAGIATGKNGPTGPRGPRGHSBPG 505
 QY 87 --GPAGTPGPRGCSVP-----PRSAFSAKRSESRVPP 117
 DB 506 LPGPGRPGPRGPGVGMVMPGPFYKACQRPRLSGMPLVSNHGVGMVSAFTVILSKAY--P 563
 QY 118 PSDAPLPEDRVLVNQGHYDAVTGFTQGVGVVYFAVHATVYRASLOFDLVKNG-ESIA 176
 DB 564 AVGAPITPDETLNRQHTDPRSGITFTCKIPGITYFYFHHKTHAWGLKNGTPVWYTYDEVYIKGLDQA--SGSAVI 623
 QY 177 SFQPRFGMKPKASISGGMVLPEDQVWVQGVGDYIGIYASIKTSTFGSFLV 232
 DB 624 TYDEVYIKGLDQA--SGSAVIMETLNDQVWVWQLPNAAGSNGLYSPRYVHSSFGSFLV 677

RESULT 8
 C1HUC
 complement subcomponent C1q chain C precursor - human
 N:Alternate names: complement subcomponent C1q gamma chain
 C:Species: Homo sapiens (man)
 C:Date: 22-May-1991 #sequence_revision 31-May-1996 #text_change 22-May-1998
 C:Accession: S14351; A03207
 R:Sejlar, G.C.; Blake, D.J.; Reid, K.B.M.
 Biochem. J. 274, 481-490, 1991
 A:Title: Characterization and organization of the genes encoding the A-, B- and C-ch
 A:Reference number: S14350; MUID:91174759; PMID:1706597
 A:Accession: S14351
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-245 <SEL>
 R:Reid, K.B.M.
 Biochem. J. 179, 367-371, 1979
 A:Title: Complete amino acid sequences of the three collagen-like regions present in
 A:Reference number: A0304; MUID:80020137; PMID:486087
 A:Accession: A03207
 A:Molecule type: protein
 A:Residues: 29-56, 'P', 58-65, 'K', 67-71, 'P', 73-83, 'K', 85-86, 'D', 88-89, 'N', 91-122 <REI>
 C:Comment: The first component of complement is a calcium-dependent complex of the c
 ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of compleme
 C:Comment: The C1q subcomponent is composed of nine subunits, six of which are disul
 dimers of the C chain. Equimolar amounts of the A, B, and C chains are found after

C:Genetics:
 A:Gene: GDB:C1QG
 A:Cross-references: GDB:128132; OMIM:120575
 A:Map position: 1p36.3-1p34.1
 A:Introns: 60/3
 C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal
 C:Keywords: complement pathway; glycoprotein; homodimer; hydroxyllysine; hydroxyproli
 F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-245/Product: complement subcomponent C1q chain B #status predicted <MAT>
 F:31-114/Domain: collagenous, triple helix <COL>
 F:121-244/Domain: complement C1q carboxyl-terminal homology <C1Q>
 F:32/Dsulfide bonds: interchain #status experimental
 F:35,39,42,45,54,61,93,96,99,105/Modified site: 4-hydroxyproline (Pro) #status expert
 F:57,72,75/Modified site: 5-hydroxylysine (Lys) #status experimental
 F:75/Binding site: carbohydrate (lys) (covalent) #status experimental

Query Match 30.1%; Score 399; DB 1; Length 245;
 Best Local Similarity 39.8%; Pred. No. 4.3e-21;
 Matches 100; Conservative 33; Mismatches 84; Indels 34; Gaps 10;

1 MRPLVLLGLAGSPPLDNKIPSLC---PGRGLPGRTHHSGSLPRDRDRDG 57
 :::::::::::-----GQANTGCGYIPGMPGLPGAGKXGQYDGLPGKGEPIPA 64
 Db 13 LKLLLLLLLR-----
 Qy 58 APGAPGEGEGRGLPGPRGDPGRBAGPAGTGPAGECVPP-----RSAFSA 108
 65 IPGRGPKGKGEGGLGHPGKNGPMPGPGMGI PGRGEGRYKQKQSVFTV 124
 Db 109 KRSESRVPPSDAPLPDRVLVNEQHYDAVTGKTCQVPGVYFAVHATVYRASLOFDL 168
 125 TR-QTHPPAPNSIRFNAVLTPGQDVTSTGKFTCKPGLYFYVYHAS-HTANLCLVL 182
 Qy 169 VKNGESIASPFGFGCKMPKASL-SGAMVLEPEDQVWQVGVGY---IGYASIKTD 224
 183 YRSGVAKVLT---FCGHTSKTNQVNSGVLLRLQVEEVM--LAVNDYDMWIGGS---D 234
 Db 225 STEGFLVYSD 235
 Qy 235 SVFSGFLRPD 245

RESULT 9

A34246
 collagen alpha 1(VIII) chain precursor - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A34246
 R:Yamauchi, N.; Benya, P.D.; van der Rest, M.; Nihomiya, Y.
 J. Biol. Chem. 264, 16022-16029, 1989
 A:Title: The cloning and sequencing of alpha1(VIII) collagen cDNAs demonstrate that type
 omains similar to those of type X collagen.
 A:Reference number: A34246; MUID:85380199; PMID:2476437
 A:Accession: A34246
 A:Molecule type: mRNA
 A:Residues: 1-744 <YAM>
 A:Cross-references: GB:J05042; NID:G164895; PIDN:AAA31204.1; PID:G164896
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
 F:21-117/Region: amino-terminal nonhelical
 F:118-571/Region: interrupted helical
 F:572-744/Region: carboxyl-terminal nonhelical
 F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 28.2%; Score 374; DB 1; Length 744;
 Best Local Similarity 34.4%; Pred. No. 7.6e-19;
 Matches 93; Conservative 36; Mismatches 79; Indels 62; Gaps 8;

24 IPSLC--PGRGLPGRTHHSGSLPRDRDRDGAPGAPGEGGRPLP----- 74
 :::::::::::-----GQANTGCGYIPGMPGLPGAGKXGQYDGLPGKGEPIPA 64
 Db 475 VPGILGPKGEGPGLDGLGPPGPGITGSPSGIPGPGKGEPELPDPGPPGVGK 534
 Qy 75 -----GPRGDP---GPRG---AGPAGTPGAGECVPPRS----- 104
 535 PGVAGLHGPGRKRALGPGQGLGPPGPPGPPGPPAPVMPPTAPAGSEYLPDMGLGIDG 594
 Db 105 -----AFSAKRSSESRV-----PSDAPLPDRVLVNEQHYDAVTGKFT 144
 145 VKTPHAAVAAKKKNGKGPAYEMPAPFTALTAFFPGVAPGAIKDRRLYNRQVYNPQTGIET 654
 Qy 145 QVWVGVGVYFAVHATVYRASLOFDLVKNGESIA-SFPGFGCKMPKASLSGAMVLEPED 203

Db 655 CEVPGVYFAVHATVYRASLOFDLVKNGESIA-SFPGFGCKMPKASLSGAMVLEPED 712
 Qy 204 QVWVGVGVYFAVHATVYRASLOFDLVKNGESIA-SFPGFGCKMPKASLSGAMVLEPED 203
 Db 713 RVFLQNPSEQAAGLYACQYVHSSRSGLY 742

RESULT 10

S23298
 collagen alpha 1(VIII) chain - chicken
 C:Species: Gallus gallus (chicken)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S23298
 R:Nihomiya, Y.; Castagna, P.; Gerecke, D.; Gordon, M.K.; Jancenko, O.; LuValle, P.;
 maguchi, N.; Olsen, B.R.
 in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic
 A:Title: The molecular biology of collagens with short triple-helical domains.
 A:Reference number: S22243
 A:Accession: S23298
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-744 <NIN>
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homol
 F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 27.9%; Score 370; DB 1; Length 744;
 Best Local Similarity 34.4%; Pred. No. 1.4e-18;
 Matches 93; Conservative 35; Mismatches 80; Indels 62; Gaps 8;

24 IPSLC--PGRGLPGRTHHSGSLPRDRDRDGAPGAPGEGGRPLP----- 74
 :::::::::::-----GQANTGCGYIPGMPGLPGAGKXGQYDGLPGKGEPIPA 64
 Db 475 VPGILGPKGEGPGLDGLGPPGPGITGSPSGIPGPGKGEPELPDPGPPGVGK 534
 Qy 75 -----GPRGDP---GPRG---AGPAGTPGAGECVPPRS----- 104
 535 PGVAGLHGPGRKRALGPGQGLGPPGPPGPPGPPAPVMPPTAPAGSEYLPDMGLGIDG 594
 Db 105 -----AFSAKRSSESRV-----PSDAPLPDRVLVNEQHYDAVTGKFT 144
 145 QVWVGVGVYFAVHATVYRASLOFDLVKNGESIA-SFPGFGCKMPKASLSGAMVLEPED 203
 655 CEVPGVYFAVHATVYRASLOFDLVKNGESIA-SFPGFGCKMPKASLSGAMVLEPED 712
 Qy 204 QVWVGVGVYFAVHATVYRASLOFDLVKNGESIA-SFPGFGCKMPKASLSGAMVLEPED 203
 Db 713 RVFLQNPSEQAAGLYACQYVHSSRSGLY 742

RESULT 11

S23779
 collagen alpha 1(VIII) chain - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: S23779
 R:Muragaki, Y.; Shiohara, C.; Inoue, M.; Ooshima, A.; Olsen, B.R.; Nihomiya, Y.
 Eur. J. Biochem. 207, 895-902, 1992
 A:Title: Alpha-1(VIII)-collagen gene transcripts encode a short-chain collagen poly
 A:Reference number: S23779; MUID:92362626; PMID:1499564
 A:Accession: S23779
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-743 <MR>
 A:Cross-references: EMBL:X66976; NID:950493; PIDN:CAA47387.1; PID:G1359953
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homol
 F:616-742/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 27.4%; Score 363.5; DB 1; Length 743;
 Best Local Similarity 32.8%; Pred. No. 4.1e-18;
 Matches 90; Conservative 35; Mismatches 78; Indels 71; Gaps 7;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 09:45:59 ; Search time 36 Seconds
(without alignments)

279.965 Million cell updates/sec

Title: US-09-944-403-42

Perfect score: 1325

Sequence: 1 MRPLVLLGLAGSPPLD.....DSTFGFLVYSDMHSPEFA 243

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1325	100.0	243	1	COT5_HUMAN
2	433.5	32.7	247	1	Q9BXJ0 homo sapien
3	424	32.0	244	1	APM1_MOUSE
4	418.5	31.6	680	1	APM1_HUMAN
5	417.5	31.5	674	1	CA1A_HUMAN
6	412	31.1	246	1	CA1A_CHICK
7	411.5	31.0	674	1	CA1A_BOVIN
8	410.5	31.0	289	1	COT7_HUMAN
9	408	30.8	680	1	CA1A_MOUSE
10	403	30.4	285	1	COT2_HUMAN
11	397	30.0	245	1	C1QC_HUMAN
12	374	28.2	744	1	CA18_RABIT
13	373	28.2	419	1	COLE_LEPMA
14	368	27.8	744	1	CA18_HUMAN
15	366.5	27.7	743	1	CA18_MOUSE
16	363	27.4	635	1	CA28_HUMAN
17	362	27.3	251	1	C1OB_HUMAN
18	360.5	27.2	253	1	C1OB_RAT
19	350	26.4	253	1	C1OB_MOUSE
20	314	23.7	245	1	C1OA_HUMAN
21	314	23.7	245	1	C1OA_MOUSE
22	298.5	22.5	255	1	GLIC_MOUSE
23	285.5	21.5	258	1	C1RF_HUMAN
24	282	21.3	215	1	C1RF_MOUSE
25	280.5	21.2	215	1	HP25_TAMSI
26	277.5	20.9	215	1	HP27_TAMSI
27	264.5	20.0	246	1	COT3_HUMAN
28	261.5	19.7	196	1	HP20_TAMSI
29	230.5	17.4	1049	1	CA13_BOVIN
30	229.5	17.3	281	1	COT1_HUMAN
31	228	17.2	684	1	CA39_HUMAN
32	224	16.9	636	1	CA13_RAT
33	223	16.8	1464	1	CA13_MOUSE

34	222.5	16.8	1758	1	CA24_CAEEL	P17140 caenorhabdi
35	221.5	16.7	170	1	CA28_MOUSE	P25318 mus musc
36	221	16.7	1019	1	CA16_CHICK	P20785 gallus gall
37	221	16.7	1262	1	CA13_CHICK	P12105 gallus gall
38	220	16.6	248	1	PSPA_CANFA	P06908 canis fami
39	220	16.6	360	1	CCD2_CAEEL	P35799 caenorhabdi
40	219.5	16.6	247	1	PSPA_RABIT	P12842 oryctolagus
41	219.5	16.6	1466	1	CA13_HUMAN	P02461 homo sapien
42	219	16.5	671	1	CA11_RAT	P02454 rattus norv
43	218.5	16.5	1453	1	CA11_MOUSE	P11087 mus musc
44	217.5	16.4	1516	1	CA1H_HUMAN	P39060 homo sapien
45	217	16.4	747	1	CA12_BOVIN	P02459 bos taurus

ALIGNMENTS

RESULT 1
ID COT5_HUMAN STANDARD; PRT; 243 AA.
AC Q9BXJ0: Q9BXJ0; 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P.O., Humes J.M.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF329841; AAK17965.1; -
CC EMBL; AL110261; CAB53702.1; -
CC Genew; HGNC:14344; C1QTNF5.
DR InterPro: IPR01073; C1Q.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; C1q; 1.
DR Pfam: PF01391; Collagen; 1.
DR PRINTS: PRO0007; COMPLEMENTC1Q.
DR SMART: SM00110; C1Q; 1.
DR PROSITE: PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 15
FT CHAIN 16 243
FT DOMAIN 30 95
FT DOMAIN 97 243
FT SEQUENCE 243 AA; 25298 MW; 7CCDA65CD47EB784 CRC64;
Query Match 100.0%; Score 1325; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.7e-85;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

FT MOD_RES 65 65 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 79 79 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 98 98 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 107 107 HYDROXYLATION (BY SIMILARITY).
FT VARIANT 113 113 M -> V.
FT CONFLICT 50 50 P -> S (IN REF. 2).
FT CONFLICT 74 74 A -> S (IN REF. 2).
FT CONFLICT 117 117 A -> G (IN REF. 2).
FT CONFLICT 148 148 G -> N (IN REF. 2).
FT CONFLICT 243 243 Y -> F (IN REF. 2).
SQ SEQUENCE 247 AA; 26841 MW; 137B687D873986C4 CRC64;

Query Match 32.7%; Score 433.5; DB 1; Length 247;
Best Local Similarity 40.6%; Pred. No. 1,2e-23;
Matches 102; Conservative 32; Mismatches 92; Indels 25; Gaps 8;

QY 1 MRPLVLALL-GLAAGSPPLDNRKPSLCPGHPG-----LPCTPGHSSQGLPGRDGDG 54
DB 4 LQALLFLILPSHADDDVTTBELAPALVPPKGCACMMAGIRPHGNHGRDGRD- 62
QY 55 RDGAPGAPGEGEGRPGIIPGRPD-----RPRGEGAPGAPGPGAGCSVPPSA 105
DB 63 -----GTGEGKEGKGDAGLGPKEGDMGATGAGGPRGPGTGRKSGPGEAAVYRSA 117
QY 106 PSKASESRVPPSPAPLPFDRVLVNEGHTDVAVTGKTCOVPRGVYAVNAHTVYRASLQ 165
DB 118 FSV-GLFTRVTP-NVPIRFTKI FYNQOHYDGSFGKFCNIRGLYTSYHTVVMKQVK 175
QY 166 PLYVNGESIASFFQFGGMPKASISGAGVRLPEPDVQVQV-GVGDYIGIVASIKTD 224
DB 176 VSLFKDKRAVLFTYQYGF-KVNDQASGSVLAHLERGVQVWQVYGGDGHNLVYDNDVND 234
QY 225 STFGSGFLVYSD 235
DB 235 STFGSLVYHD 245

RESULT 3
APM1 HUMAN STANDARD; PRT; 244 AA.
ID APM1 HUMAN
AC 015648;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (Adipose most abundant gene transcript 1) (apm-1) (Gelatin-
DE binding protein).
GN APM1 OR ACRP30 OR GBP28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=9622411; PubMed=8619847;
RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA Matsubara K.;
RT "cDNA cloning and expression of a novel adipose specific collagen-like
RT factor, apm1 (adipose most abundant gene transcript 1).";
RL Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9916984; PubMed=10095105;
RA Saito K., Tobe T., Minoshima S., Asakawa S., Sumiya J., Yoda M.,
RA Nakano Y., Shimizu N., Tomita M.;
RT "Organization of the gene for gelatin-binding protein (GBP28).";
RL Gene 229:67-73(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99333693; PubMed=10403784;
RA Schaeffler A., Orso B., Palitzsch K.D., Buechler C., Drobnik W.,
RA Fuerst A., Scholmerich J., Schmitz G.;

```

```

RT "The human apm-1, an adipocyte-specific gene linked to the family of
RT TNF's and to genes expressed in activated T cells, is mapped to
RT chromosome 1q21.3-q23, a susceptibility locus identified for familial
RT combined hyperlipidemia (FCH).";
RL Biochem. Biophys. Res. Commun. 260:416-425(1999).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20417747; PubMed=10961870;
RA Yokota T., Ohtani K., Takahashi I., Ishikawa J., Matsuyama A.,
RA Ouchi N., Kihara S., Funahashi T., Tamer A.J., Tomiyama Y.,
RA Matsuzawa Y.;
RT "Adiponectin, a new member of the family of soluble defense collagens,
RT negatively regulates the growth of myelomonocytic progenitors and the
RT functions of macrophages.";
RL Blood 96:1723-1732(2000).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=20440368; PubMed=10982546;
RA Ouchi N., Kihara S., Arita Y., Okamoto Y., Maeda K., Kuriyama H.,
RA Hotta K., Nishida M., Takahashi M., Muraguchi M., Ohmoto Y.,
RA Nakamura T., Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Adiponectin, an adipocyte-derived plasma protein, inhibits
RT endothelial NF-kappaB signaling through a C/EBP-dependent pathway.";
RL Circulation 102:1296-1301(2000).
RN [6]
RP FUNCTION.
RX MEDLINE=21372498; PubMed=11479627;
RA Yamauchi T., Kamon J., Maki H., Terauchi Y., Kubota N., Hara K.,
RA Mori Y., Ide T., Murakami K., Tsudoyama-Kasaoaka N., Ezaki O.,
RA Akanuma Y., Gavrilova O., Vinson C., Retman M.L., Kagechika H.,
RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA Froguel P., Kadowaki T.;
RT "The fat-derived hormone adiponectin reverses insulin resistance
RT associated with both lipodystrophy and obesity.";
RL Nat. Med. 7:941-946(2001).
RN [7]
RP VARIANT ADIPONECTIN DEFICIENCY CYS-112.
RX MEDLINE=20378830; PubMed=10918532;
RA Takahashi M., Arita Y., Yamagata K., Matsukawa Y., Okutomi K.,
RA Horie M., Shimomura I., Hotta K., Kuriyama H., Kihara S., Nakamura T.,
RA Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Genomic structure and mutations in adipose-specific gene,
RT adiponectin.";
RL Int. J. Obes. Relat. Metab. Disord. 24:861-868(2000).
RN [8]
RP VARIANTS ARG-84; MET-117; THR-164; SER-221 AND PRO-241.
RX MEDLINE=21671103; PubMed=11812766;
RA Hara K., Boutin P., Mori Y., Tobe K., Dina C., Yasuda K., Yamauchi T.,
RA Otake S., Okada T., Eto K., Kadowaki H., Hagura R., Akanuma Y.,
RA Yazaki Y., Nagai R., Tanigawa M., Matsubara K., Yoda M., Nakano Y.,
RA Kimura S., Tomita M., Kimura S., Ito C., Froguel P., Kadowaki T.;
RT "Genetic variation in the gene encoding adiponectin is associated with
RT an increased risk of type 2 diabetes in the Japanese population.";
RL Diabetes 51:536-540(2002).
RN [9]
RP FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
RN SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
RN ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPA-B SIGNALING
RN THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
RN EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
RN CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
RN [10]
RP SUBUNIT: HOMODIGLIMER (POTENTIAL).
RN [11]
RP TISSUE SPECIFICITY: EXPRESSED ONLY IN THE ADIPOSE TISSUE.
RN [12]
RP DISASS: Defects in APM1 are the cause of adiponectin deficiency,
RN resulting in very low concentration of plasma adiponectin.
RN [13]
RP insulin resistance, and diabetes type 2.
RN [14]
RP PHARMACEUTICAL: Adiponectin might be used in the treatment of
RN diabetes type 2 and insulin resistance.
RN [15]
RP SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
RN [16]
RP SIMILARITY: CONTAINS 1 C1Q DOMAIN.
RN [17]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D45371; BAA08227.1; -
 DR EMBL: AB012165; BAA86716.1; JOINED.
 DR EMBL: AB012164; BAA86716.1; JOINED.
 DR EMBL: AJ131460; CAB52413.1; -
 DR EMBL: AJ131461; CAB52413.1; JOINED.
 DR MIM: 605441; -
 DR InterPro: IPR001073; C1Q.
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF00386; C1q; 1.
 DR Pfam: PF01391; Collagen; 1.
 DR PRINTS: PR00007; COMPLEMENTC1Q.
 DR SMART: SM00110; C1Q; 1.
 DR PROSITE: PS01113; C1Q; 1.
 KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
 KM Polymorphism; Disease mutation; Obesity; Diabetes mellitus.
 FT SIGNAL 1 14
 FT CHAIN 15 244
 FT DOMAIN 42 107
 FT DOMAIN 108 244
 FT DISULFID 36 36
 FT MOD_RES 44 44
 FT MOD_RES 47 47
 FT MOD_RES 53 53
 FT MOD_RES 62 62
 FT MOD_RES 71 71
 FT MOD_RES 76 76
 FT MOD_RES 86 86
 FT MOD_RES 95 95
 FT MOD_RES 104 104
 FT VARIANT 84 84
 FT VARIANT 112 112
 FT VARIANT 117 117
 FT VARIANT 164 164
 FT VARIANT 221 221
 FT VARIANT 241 241
 FT SEQUENCE 244 AA; 26414 MW; 64D8C6C1204B1018 CRC64;
 SQ
 Query Match 32.0%; Score 424; DB 1; Length 244;
 Best Local Similarity 40.5%; Pred. No. 5.3e-23;
 Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;

QY 6 VLLILGLAAGSPRLDNNKIPSLCRPH-----PG-----LPGTGHHGSGQ 44
 DB 7 VLLILAL-----PGHDETTTGGVLLPLPKGACGMMAGIPIHRCNH 50
 QY 45 GLPGRDGRDGRDGAFAPEKGEKGRPGILPGRPD-----PGPRGEAGPAAPTGA 95
 DB 51 GAFGRDGRD-----GTRGEKGEKDPGLIGPKGIGETGVPGAGPRPGIQRKGEF 104
 QY 96 GECVPPPSRAFSAKSESSEKVRPPSPALPFRVLVNEGQHDVATGKTCTCOVPGVYRAY 155
 DB 105 GEGAYVYRSAPFSV-GLETVVITP-NMIRFTKIFYNQDNHVDGSGKFKHCNIPGLYRAY 162
 QY 156 HATVYRASLQPLVNVNGES-IASPFQFGWPKPSSLGCAVNRLEPEDQWVQY-GVGD 213
 DB 163 HTTVYKDVKVSLEPKDKAMLFTYDQYQENNVDA--SGSVLIHLVEGDQWLVQYGGEGE 220
 QY 214 YIGIYASIKTDSIFSGFLVYSD 235
 DB 221 RNLGLVADNDNDSTFTGFLLYHD 242

RESULT 4
 ID CAIA HUMAN STANDARD; PRT; 680 AA.
 AC 003692;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Collagen alpha 1(X) chain precursor.
 GN COL10A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Butheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxId=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92109659; PubMed=1764025;
 RA Thomas J.T., Cresswell C.J., Rash B., Nicolai H., Jones T.,
 RA Solomon E., Grant M.E., Boot-Handford R.P.;
 RT "The human collagen X gene. Complete primary translated sequence and
 RT chromosomal localization.";
 RL Biochem. J. 280:617-623 (1991).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93012005; PubMed=1397333;
 RA Reichdeberger E., Beier F., Luvallie P., Olsen B.R., von der Mark K.,
 RA Bertling W.M.;
 RT "Genomic organization and full-length cDNA sequence of human collagen
 RT X.";
 RL FEBS Lett. 311:305-310 (1992).
 RN (3)
 RP SEQUENCE FROM N.A.
 RA Beier F., Lamm M.B., von der Mark K.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A.
 RA Williams S.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN (5)
 RP SEQUENCE OF 52-680 FROM N.A.
 RX MEDLINE=92267014; PubMed=1587271;
 RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
 RT "Cloning of the human and mouse type X collagen genes and mapping of
 RT the mouse type X collagen gene to chromosome 10.";
 RL Eur. J. Biochem. 206:217-224 (1992).
 RN (6)
 RP SEQUENCE OF 561-666 FROM N.A.
 RX MEDLINE=91243838; PubMed=2037056;
 RA Apte S., Mattei M.-G., Olsen B.R.;
 RT "Cloning of human alpha 1(X) collagen DNA and localization of the
 RT COL10A1 gene to the q21-q22 region of human chromosome 6.";
 RL FEBS Lett. 282:393-396 (1991).
 RN (7)
 RP SEQUENCE OF 547-655 FROM N.A.
 RX MEDLINE=92077285; PubMed=1743401;
 RA Reichdeberger E., Aigner T., von der Mark K., Stoeck H., Bertling W.;
 RT "In situ hybridization studies on the expression of type X collagen
 RT in fetal human cartilage.";
 RL Dev. Biol. 148:562-572 (1991).
 RN (8)
 RP REVIEW ON VARIANTS.
 RX MEDLINE=97255959; PubMed=9101290;
 RA Kuivaniemi H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 RT associated collagen (type IX), and network-forming collagen (type X)
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315 (1997).
 RN (9)
 RP VARIANTS SMCD ASP-598 AND PRO-614.
 RX MEDLINE=94136476; PubMed=8104336;
 RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
 RA Grant M.E., Boot-Handford R.P.;

RT "Aminoacid substitutions of conserved residues in the
RT [alpha 1(I) collagen chain of type X collagen
RT occur in two unrelated families with metaphyseal chondrodysplasia
RT type Schmid." ;
RN Am. J. Hum. Genet. 54:169-178(1994) .
RN [10]
RN VARIANT SMCD ARG-591.
RN MEDLINE=94227470; PubMed=80040099;
RX McIntosh I., Abbott M.H., Watanabe M.L., Olsen B.R., Francomano C.A.;
RT "Additional mutations of type X collagen confirm COL10A1 as the
RT Schmid metaphyseal chondrodysplasia locus." ;
RN Hum. Mol. Genet. 3:303-307(1994) .
RN [11]
RN VARIANT SMCD VAL-618.
RX MEDLINE=95181449; PubMed=7876225;
RX Chan D., Cole W.G., Rogers J.G., Bateman J.F. ;
RT "Type X collagen multimer assembly in vitro is prevented by a Gly618
RT to Val mutation in the alpha 1(X) NC1 domain resulting in Schmid
RT metaphyseal chondrodysplasia." ;
RN J. Biol. Chem. 270:4558-4562(1995) .
RN [12]
RN VARIANT SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648
RN MEDLINE=95331767; PubMed=7607655;
RX Bonaventure J., Chamande P.F., Maroteaux P. ;
RA "Mutations in three subdomains of the carboxy-terminal region of
RT collagen type X account for most of the Schmid metaphyseal
RT dysplasia." ;
RN Hum. Genet. 96:58-64(1995) .
RN [13]
RN VARIANT SMCD PRO-600.
RX MEDLINE=96375754; PubMed=8782043;
RX Wallis G.A., Raeb J., Sykes B., Bonaventure J., Maroteaux P.,
RA Zabel B., Wynne-Davies R., Grant M.E., Boot-Handford R.P. ;
RT "Mutations within the gene encoding the alpha 1 (X) chain of type X
RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
RT not several other forms of metaphyseal chondrodysplasia." ;
RN J. Med. Genet. 33:450-457(1996) .
RN [14]
RN VARIANT SMCD GLU-18 AND ARG-18.
RX MEDLINE=97220591; PubMed=9067753;
RX Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y. ;
RA "Mutations in the N-terminal globular domain of the type X collagen
RT gene (COL10A1) in patients with Schmid metaphyseal
RT chondrodysplasia." ;
RN Hum. Mutat. 9:131-135(1997) .
RN [15]
RN VARIANT SMCD GLU-595.
RX MEDLINE=99057503; PubMed=9837818;
RX Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H. ,
RA Nakamura Y. ;
RA "Mutation of the type X collagen gene 'COL10A1' causes
RT spondylometaphyseal dysplasia." ;
RN Am. J. Hum. Genet. 63:1659-1662(1998) .
RN [16]
RN VARIANT SMCD CYS-597.
RX MEDLINE=99069781; PubMed=9852679;
RX Sawai H., Ida A., Nakata Y., Koyama K. ;
RT "Novel missense mutation resulting in the substitution of tyrosine by
RT cysteine at codon 597 of the type X collagen gene associated with
RT Schmid metaphyseal chondrodysplasia." ;
RN J. Hum. Genet. 43:259-261(1998) .
RN [17]
RN FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC
CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
CC [1]
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- PPM: PROLINES AT THE THIRD POSITION OF THE TRIPLET REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SCHMID TYPE
CC METAPHYSEAL CHONDRODYSPLASIA (SMCD), A DOMINANTLY INHERITED
CC DISORDER OF THE OSSEOUS SKELETON. THE CARDINAL FEATURES OF THE
CC PHENOTYPE ARE MILD SHORT STATURE, COXA VARA AND A WADDLING GAIT.
CC RADIOGRAPHY USUALLY SHOWS SCLEROSIS OF THE RIBS, FLARING OF THE
CC METAPHYSESES, AND A WIDE IRREGULAR GROWTH PLATE, ESPECIALLY OF THE

```
CC KNEES. DEFECTS IN COL10A1 ARE THE CAUSE OF SPONDYLOMETAPHYSAL  
CC DYSPLASIA (SMD). SMD COMPRISES A HETEROGENEOUS GROUP OF HERITABLE  
CC SKELETAL DYSPLASIAS CHARACTERIZED BY MODIFICATIONS OF THE  
CC VERTEBRAL BODIES OF THE SPINE AND METAPHASES OF THE TUBULAR BONES.  
CC -1- SIMILARITY: STRONG; TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.  
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.ebi.ch/announcement/  
CC or send an email to license@ebi.ac.uk).  
CC -----  
DR EMBL; X60382; CAA42933.1; -;  
DR EMBL; X65120; CAA46236.1; -;  
DR EMBL; X98568; CAA67178.1; -;  
DR EMBL; AL121963; CAB87590.1; -;  
DR EMBL; S68531; AAC60615.1; -;  
DR EMBL; X58879; CAA41686.1; -;  
DR EMBL; M74050; AAA61221.1; -;  
DR EMBL; X72579; CAA51170.1; -;  
DR EMBL; X72580; CAA51170.1; JOINED.  
DR PIR; S15826; S15826.  
DR PIR; S30086; S30086.  
DR PIR; AA3901; AA3901.  
DR PIR; S18249; S18249.  
DR PIR; S21856; S21856.  
DR PIR; S26396; S26396.  
DR Genew; HGNC:2185; COL10A1.  
DR MIM; 120110; -;  
DR MIM; 156500; -;  
DR InterPro; IPR001073; C1q.  
DR InterPro; IPR000087; Collagen.  
DR Pfam; PF00386; C1q; 1.  
DR PRINTS; PR00007; COMPLEMENTC1Q.  
DR SMART; SM00110; C1Q; 1.  
DR PROSITE; PS0113; C1Q; 1.  
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
KW Cartilage; Collagen; Signal; Disease mutation; Polymorphism.  
KM SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 680 COLLAGEN ALPHA 1(X) CHAIN.  
FT DOMAIN 19 56 NONHELIICAL REGION (NC2).  
FT DOMAIN 57 519 TRIPLE-HELICAL REGION.  
FT DOMAIN 520 680 NONHELIICAL REGION (NC1).  
FT DOMAIN 545 680 C1Q.  
FT VARIANT 18 18 G->E (IN SMD).  
FT FTID=VAR_001838.  
FT FTID=VAR_001839.  
FT VARIANT 18 18 G->R (IN SMD).  
FT VARIANT 545 545 /FTID=VAR_001840.  
FT VARIANT 591 591 /FTID=VAR_001841.  
Query Match 31.6%; Score 418.5; DB 1; Length 680;  
Best Local Similarity 34.9%; Pred. No. 3.5e-22;  
Matches 98; Conservative 34; Mismatches 82; Indels 67; Gaps 6  
  
Oy 15 GSPPIDDNKIISLCGHPGLPCTPDHNSOGLPGHRDGRDYGAGARGEKSGGRPGIP 74  
Db 401 GNPGLPGRKDPGVGGRRPGRLCPORVPRAAKMGPHNGSRGAGRGAGIGITGTGRLPGRIP 460  
Oy 75 -----GPRGDGFPERGEA-----GPAGPTGAAGCS 99  
Db 461 GPFGSKDPSGSPGPRPAGIATKIANGTPRGVGRGPRGSGSEKGPLPGRPPGPPGCAV 520  
Oy 100 VP-----PPSAFAKSSESIVPPRSADAPLYPFRDLVNE 132
```

```

Dr Db 521 MPEGRKAGQBPSSGPFVLVSANQGVGMPSAFVTLISKAY--PAIGTPIPDKILYNR 578
Oy 133 QGHATDVAATGKTCQVPGYTYTAVAHATYTRASLDGPDLYKNESIA-SFFQFGGPKPKASL 191
Db 579 QQHDDPRGRTGTCQIPGITYPSHHVHWGTHWVGLYKNGGTPWVATYDEYTKGLDGA-- 636
Oy 192 SGAMVRLRPEDQVWVGVDYGIYASIKTDSFSGFLV 232
Db 637 SGSAIIDLTDENDQVWQLDPNAMESGLVSSSEYVSHSFGSLV 677

RESULT 5
ID CA1A_CHICK STANDARD; PRT; 674 AA.
AC P08125;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN COL10A1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE OF 48-674 FROM N.A., AND SEQUENCE OF 103-117 AND 453-466.
RX MEDLINE=86168277; PubMed=3082876;
RA Nimomiya Y., Gordon M., van der Rest M., Schmid T., Linsemayer T.,
RA Olsen B.R.;
RT "The developmentally regulated type X collagen gene contains a long
RT open reading frame without introns.";
RL J. Biol. Chem. 261:5041-5050(1986).
RN [2]
RP SEQUENCE OF 1-75 FROM N.A.
RX MEDLINE=89054019; PubMed=2461368;
RA Luvalle P., Nimomiya Y., Rosenblum N.D., Olsen B.R.;
RT "The type X collagen gene. Intron sequences split the 5'-untranslated
RT region and separate the coding regions for the non-collagenous amino-
RT terminal and triple-helical domains.";
RL J. Biol. Chem. 263:18378-18385(1988).
RN [3]
RP REVISIONS TO C-TERMINUS.
RX MEDLINE=89380109; PubMed=2476437;
RA Yamaguchi N., Benay P.D., van der Rest M., Nimomiya Y.;
RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
RT demonstrate that type VIII collagen is a short chain collagen and
RT contains triple-helical and carboxyl-terminal non-triple-helical
RT domains similar to those of type X collagen.";
RL J. Biol. Chem. 264:16022-16029(1989).
RN -1-
RP FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC
RP CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
RP MINERALIZATION ZONES OF HYALINE CARTILAGE.
RN -1-
RP SUBUNIT: HOMOTRIMER.
RN -1-
RP PIV: PROLINES ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
RN -1-
RP SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
RN -1-
RP SIMILARITY: CONTAINS 1 C1Q DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement. See http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sb.ch.
DR EMBL: M13496; AAA48736.1; ALY_SEQ.
DR EMBL: J04194; AAA48634.1; -.
DR PIR: A31896; A31896.
DR InterPro: IPR001073; C1q.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; C1q; 1.

```

```

DR PFAM: PF01391; Collagen: 8.
DR PRINTS: PR00007; COMPLEMENTC1Q.
DR SMART, SM00110; C1Q; 1.
DR EXTRACELL; PS01113; C1Q; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cartilage; Collagen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 674 COLLAGEN ALPHA 1(X) CHAIN.
FT DOMAIN 19 52 NONHELICAL REGION (NC).
FT DOMAIN 53 512 TRIPLE-HELICAL REGION.
FT DOMAIN 513 674 NONHELICAL REGION (NCL).
FT DOMAIN 539 674 C1Q.
FT MOD RES 453 453 HYDROXYLATION.
FT MOD RES 456 456 HYDROXYLATION.
SQ SEQUENCE 674 AA; 66434 MW; EAB48B1EF174B145 CRC64;

Query Match 31.5%; Score 417; DB 1; Length 674;
Best Local Similarity 36.0%; Pred. No. 4,4e-22;
Matches 96; Conservative 34; Mismatches 84; Indels 53; Gaps 6

QY 14 AGSPPLDNDKPLPSLCRPHGLPRTGTHGHSQGLPGRDGDGDRGAPGARGEGEGRPGI 73
DB AGHPPLPGPVGPQGVKGVPGIENGEPGRKPSGIPGVKRPITGPRMFGARGANGAGAPGL 467
QY 74 PGPR-----GDPGRGEA-----GPAPTGAEGCVP----- 101
DB PGAPGIVTKGLRGPMGPPLGPPPKGNSGRLGPPGPPGPGQSTIPGIVKSGSRRLS 527
QY 102 -----PSASFAKRSBSRVPSPDAPLPEDVLYNEQHDYAVNGKFTC 145
DB GMSFMKAGANQALGMPVSAFVLLISKV--PGATVPILKEDKILYNRQOHYPRGTGIFTC 585
QY 146 QVPGVYVAVNAVTVRASLQEPFLVNGSGISA-SFQPFQGMKPKPALSGLGAVRLEPEDQ 204
DB RHPGLYTSYNHNAKGTVMVWVALYNGSGVVMVTTYDEYQGLYDQ--SGSAVIDLMENDQ 643
QY 205 VWVQVGVGDYIGIVASIKTDSFSGFL 231
DB 644 VTLQLPNSESNGLYSSEYVHSSFSGFL 670

RESULT 6
C1QC_MOUSE STANDARD; PRT; 246 AA.
ID C1QC_MOUSE
AC Q02105;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Complement C1q subcomponent, C chain precursor.
GN C1QC OR C1QC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR, TISSUE=Macrophage;
RX MEDLINE=93011118; PubMed=1356591;
RA Petry F., Reid K.B.M., Loos M.;
RT "Isolation, sequence analysis and characterization of cDNA clones
RT coding for the C chain of mouse C1q. Sequence similarity of
RT complement subcomponent C1q, collagen type VIII and type X and
RT precerebellin.";
RL Eur. J. Biochem. 209:129-134(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c; TISSUE=Liver;
RX MEDLINE=96186528; PubMed=8606057;
RA Petry F., McClive P.J., Botto M., Morley B.J., Morahan G., Loos M.;
RT "The mouse C1q genes are clustered on chromosome 4 and show
RT conservation of gene organization.";
RL Immunogenetics 43:370-376(1996).
CC 1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD

```

Db	187	LARVASECDHMFN-SKQVS-SGCALLRLRGDEWV--LSVNDYNGMVGVEGSSNSVPSGRL	242
QY	232	VYSD 235	
	:	:	
Db	243	LFPD 246	
RESULT 7			
ID	CAIA_BOVIN	STANDARD;	PRT; 674 AA.
AC	P23206;		
DT	01-NOV-1991 (Rel. 20, Created)		
DT	01-NOV-1991 (Rel. 20, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	Collagen alpha 1(X) chain precursor.		
GN	COL10A1.		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Cartilage;		
RX	MEDLINE=9111311; PubMed=1703407;		
RA	Thomas J.T., Kwan A.P.L., Grant M.E., Boot-Handford R.P.;		
RT	"Isolation of cDNAs encoding the complete sequence of bovine type X		
RT	collagen. Evidence for the condensed nature of mammalian type X		
RT	collagen genes."		
RL	Biochem. J. 273:141-148(1991).		
CC	-1- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC		
CC	CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE		
CC	MINERALIZATION ZONES OF HYALINE CARTILAGE.		
CC	-1- SUBUNIT: HOMOTRIMER.		
CC	-1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING		
CC	UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.		
CC	-1- SIMILARITY: STRONG. TO ALPHA 1 AND 2 TYPE VII COLLAGENS.		
CC	-1- SIMILARITY: CONTRAINS 1 C1Q DOMAIN.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its use		
CC	by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X53556; CAA37624.1; -.		
DR	PIR; S13301; S13301.		
DR	InterPro; IPR001073; C1q.		
DR	InterPro; IPR000087; Collagen.		
DR	Pfam; PF00386; C1q; 1.		
DR	Pfam; PF01391; Collagen; 9.		
DR	PRINTS; PR00007; COMPLEMENTC1Q.		
DR	ProDom; PD000007; Collagen; 1.		
DR	SMART; SM00110; C1Q; 1.		
DR	PROSITE; PS01113; C1Q; 1.		
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation;		
KW	Cartilage; Collagen; Signal; Glycoprotein.		
FT	SIGML	1	18
FT	CHAIN	19	674
FT	DOMAIN	19	56
FT	DOMAIN	57	519
FT	DOMAIN	520	674
FT	DOMAIN	539	674
FT	DISULFID	194	197
FT	MOD_RES	460	460
FT	MOD_RES	463	463
FT	CANBOHYD	611	611
SEQ	SEQUENCE	674 AA;	65546 MW; CD4CA73A03E004CA CRC64;
Query Match		31.1%;	Score 411.5; DB 1; Length 674;
Best Local Similarity		34.9%;	Pred. No. 1.le-21;

Matches 96; Conservative 34; Mismatches 84; Indels 61; Gaps 6;

OY 15 GSPPLDDDKITSLCGHPGLPCTPGHNSOGLPGRDGDRDAGAPGKGEKGGPPLP 74
 Db 401 GNPGLPGKGDGPIAGSPGLPGPVPAGAKGVPGHNGEAGRGVGTCTGTGPIGPPIIP 460
 OY 75 -----GPRGDPGPRGEA-----GPAFTGPAEGCS 99
 Db 461 GPRGSKDVGTPGPPGPAIGAVKGLNGPFGPPGPPGKNAGEPGLPGPPGPPGQOVA 520
 OY 100 VP-----PSASAKSESRVPPSDAPLPDPRLVNEQGYDA 138
 Db 521 IPEDVPAKQGRPEVSANOGVTGMPVSAFTVILSKAY--PAIGTPIPPDKILYKQGHYDP 578
 OY 139 VNGKFTCCOVPGVYFVAVHATYVRASLQFDLVNKGSSIA-SFQPFQGMFKPASLSGAMV 197
 Db 579 RTGIFTCXKPGIYTSYHIVKGTAMVGLYKNGTPVMTTYDEYIKGYLDQA--SGSANVI 636
 OY 198 RLEPEDQVWVGVGVGYIGIYASIKTSTFSGFLV 232
 Db 637 DLTENDQVWLQLPNAGSNGLYSPEYHSSFSGFLV 671

RESULT 8
 COT7 HUMAN STANDARD; PRT; 289 AA.

AC 09BKJ2;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Complement-c1q tumor necrosis factor-related protein 7 precursor.
 GN C1QMF7 OR CTRP7.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.;
 RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RA Strassberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL; AF329839; AAK1963.1;
 CC EMBL; BC022187; AAK22187.1;
 CC Genew; HGNC:14342; C1QTNF7.
 DR InterPro; IPR001073; C1Q.
 DR InterPro; IPR000867; Collagen.
 DR Pfam; PF00386; C1Q; 1.
 DR Pfam; PF01391; Collagen; 2.
 DR PRINTS; PRO0007; COMPLEMENTC1Q.
 DR SMART; SM00110; C1Q; 1.
 DR PROSITE; PS01113; C1Q; 1.
 KM Collagen; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 1 289 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
 FT DOMAIN 38 139 RELATED PROTEIN 7.
 FT DOMAIN 141 276 COLLAGEN-LIKE.
 FT C1Q.

SO SEQUENCE 289 AA; 30683 MM; A61609FF86D26946 CRC64;
 Query Match 31.0%; Score 410.5; DB 1; Length 289;
 Best Local Similarity 40.0%; Pred. No. 5,4e-22;
 Matches 98; Conservative 26; Mismatches 84; Indels 37; Gaps 7;

OY 24 IPBL--CRGHPGLPGTPGHHGSOGLPGRDGRDGRDAP-----GAPGKKG 66
 Db 36 IPGLPGPPGPGANGSPGPHGRIGLPORDRDRKKEKSGKTAGLRGKTGPIGLAGEKG 95
 OY 67 EGGPRGLPGPRGDPGPRGEAGPAGPTGPAGE-----C-SVPPRGAFAKRS 111
 Db 96 DQGETKKGPIGPRGEGEVGPIGPPGKDRGQDPLPGVCRCSIVLKSAFSGIT 155
 OY 112 ESRVPPSDAPLPDRVLVNEQGYDAVTKFTCOVGVYFVAVHATYVRASLQFDLVN 171
 Db 156 TSY--PBERLPIINPKVLFNEGHTYNATGKFAFGIYFSDITLANKHLAIGLVN 213
 OY 172 GESIASFPQFGGMPKPKPASLSGAMVRLPEPDQVWVGVGVGYIGIYASIK-TDSTSGF 230
 Db 214 GQYRIKTFDANTGNHDA--SGSTVIYILQPDDEVMLEIFITDQNGLFSDFGMADSLFSGF 271

OY 231 LVYSD 235
 Db 272 LLYVD 276

RESULT 9
 CALA MOUSE STANDARD; PRT; 680 AA.

AC 005306;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Collagen alpha 1(X) chain precursor.
 GN COL10A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=93143676; PubMed=8424763;
 RA Elima K., Beroia I., Rosati R., Metsaranta M., Garofalo S., Perala M.,
 de Crombrughe B., Vuorio E.;
 RT "The mouse collagen X gene: complete nucleotide sequence, exon
 RT structure and expression pattern.";
 RL Biochem. J. 289:247-253 (1993).

RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv; TISSUE=Liver;
 RX MEDLINE=9328750; PubMed=8477738;
 RA Kong R.Y.C., Kwan K.M., Lau E.T., Thomas J.T., Boot-Handford R.P.,
 Grant M.B., Cheah K.S.E.;
 RT "Intron-exon structure, alternative use of promoter and expression of
 RT the mouse collagen X gene, Col10a-1.";
 RL Eur. J. Biochem. 213:99-111 (1993).

RN 13
 RP SEQUENCE OF 51-680 FROM N.A.
 RC STRAIN=DBA/2J;
 RX MEDLINE=92267014; PubMed=1587271;
 RA Apté S.S., Seldin M.F., Hayashi M., Olsen B.R.;
 RT "Cloning of the human and mouse type X collagen genes and mapping of
 RT the mouse type X collagen gene to chromosome 10.";
 RL Eur. J. Biochem. 206:217-224 (1992).

RN 14
 RP SEQUENCE OF 385-627 FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=92182017; PubMed=1543751;
 RA Elima K., Metsaranta M., Kallilo J., Perala M., Beroia I.,
 Garofalo S., de Crombrughe B., Vuorio E.;
 RT "Specific hybridization probes for mouse alpha 2(Ix) and alpha 1(X)

QY	DB	177	SFPQFPGWPKPASLSGAMVRLPEDDQVWVGVGDYIGYISITDSTFSGFLV	232
DB	564	AVGAPRPFDEILVNRQOHVDRPSGIFCTCKIPGIYSHVHVHVGTHVWVGLYNGNGPTWV	623	
QY	177	SFPQFPGWPKPASLSGAMVRLPEDDQVWVGVGDYIGYISITDSTFSGFLV	232	
DB	624	TYDEYSKGYLDOA--SSGAIMELTENDQVWLQLPNAESNGLYSSEYVSHSFGSLV	677	
QY	109	KRESSEKVPSPSDA-----PLPDRVLVNEQGHYDAVTKFTQ	146	

```

Dh 131 KKEPGLPGSCSGSHTSAPSVAVNTKSYPREBLPIKFDKLINMGEGHVAASGKFCG 190
Qy 147 VPGVYFVAHATVYRASLQFDLVKNGESIASFPFGGKPKPKASLSGAMRLRPBQV 206
Db 191 VPGVYFVATYITLANKHLAIGVHNNQYRIKRFEDNANTGNHVA--SGSTILALQGBE 248
Qy 207 VQVGVGDVYIGI-VYASIKTDSFGFLVYSD 235
Db 249 LQIFYSEQNGLFYDPYWTDSLFLGFLIYAD 278

RESULT 11
C1QC_HUMAN
ID C1QC_HUMAN STANDARD; PRT; 245 AA.
AC P02747; Q96DL2; Q96H05.
DT 21-JUN-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement C1q subcomponent, C chain precursor.
GN C1QC OR C1QG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Monocytes;
RX Medline=91174759; Pubmed=1706597;
RX Sellar G.C., Blake D.J., Reid K.B.M.;
RT "Characterization and organization of the genes encoding the A-, B-
RT and C-chains of human complement subcomponent C1q. The complete
RT derived amino acid sequence of human C1q.";
RL Biochem. J. 274:481-490(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX Nishi T., Nakagawa S., Senoh A., Mitunuchi H., Inagaki H., Suzuki Y.,
RX Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RX Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RX Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 29-122.
RX MEDLINE=80020137; Pubmed=466087;
RX Reid K.B.M.;
RT "Complete amino acid sequences of the three collagen-like regions
RT present in subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 179:367-371(1979).
RN [5]
RP REVIEW OF C1Q DEFICIENCY.
RX MEDLINE=98450587; Pubmed=9777412;
RX Petry F.;
RT "Molecular basis of hereditary C1q deficiency.";
RL Immunobiology 199:286-294(1998).
RN -1- FUNCTION: C1Q ASSOCIATES WITH THE PROXENYMS CIR AND C1S TO YIELD
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROXENYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
CC OF NINE SUBUNTS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC THE C CHAIN.
CC -1- PTM: O-LINKED GLYCANS CONSIST OF GLC-GAL DISACCHARIDES.

```

CC	-1- DISEASE: DEFECTS IN C10G ARE A CAUSE OF C1Q DEFICIENCY. IT IS A
CC	RARE GENETIC DISORDER WHICH IS ASSOCIATED WITH RECURRENT
CC	INFECTIONS AND A HIGH PREVALENCE OF LUPUS ERYTHEMATOSUS-LIKE
CC	SYMPTOMS. IT IS CHARACTERIZED BY A LOSS OF ACTIVATION OF THE
CC	COMPLEMENT CLASSICAL PATHWAY.
CC	-1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC	-1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC	-----
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isb-sib.ch/announce/
CC	OR SEND AN EMAIL TO license@isb-sib.ch).
CC	-----
DR	EMBL; AK057792; BAB71575.1; -
DR	EMBL; BC009016; AAH09016.1; -
DR	PIR; A03207; C1HQQC.
DR	PIR; S14351; S14351.
DR	GENEW; HGNC:1245; C1QG.
DR	MIM; 120575; -
DR	InterPro; IPR001073; C1Q.
DR	InterPro; IPR000087; Collagen.
DR	Pfam; PF00386; C1q; 1.
DR	Pfam; PF01391; Collagen; 1.
DR	PRINTS; PRO0007; COMPLEMENTC1Q.
DR	SMART; SM00110; C1Q; 1.
DR	PROSITE; PS01113; C1Q; 1.
DR	Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW	Repeat; Signal; Disease mutation.
KW	Signal; 1
FT	CHAIN 28
FT	FT 29 245
FT	FT 31 112
FT	FT 113 245
FT	FT 32 32
FT	FT 36 36
FT	FT 39 39
FT	FT 42 42
FT	FT 45 45
FT	FT 54 54
FT	FT 57 57
FT	FT 63 63
FT	FT 66 66
FT	FT 71 71
FT	FT 75 75
FT	FT 75 75
FT	FT 81 81
FT	FT 84 84
FT	FT 84 84
FT	FT 93 93
FT	FT 96 96
FT	FT 99 99
FT	FT 105 105
FT	FT 43 43
FT	FT 14 14
FT	FT 23 23
FT	FT 57 57
FT	FT 66 66
FT	FT 72 72
FT	FT 84 84
FT	FT 87 87
FT	FT 90 90
FT	FT 215 215
FT	FT 245 AA; 25774 MW; 30.0%; Score 397; DB 1; Length 245;
FT	FT 39.8%; Pred. No. 4e-21; 84; Indels 34; Gaps 10
FT	FT 100; Conservative 33; Mismatches 34; Indels 34; Gaps 10
QY	QY 1 MPRLVLLILGLAAGSPPLDDNKIPSLC---GCHPGLPPTPGHHSQGLPGRDGRDGRDQ 57

DB 13 LKLLLLLLLPLR-----GQANTGCVIGPMGLPGAPGKDGVDGLPGKGPPIA 64

QY 58 AFGAPGKGGGGRPGLPGRGDPGRGGEAPGPTGPAEGCVPP-----SAPSA 108

DB 65 IPIGIPKQKQKGPGLPGHPGKNGGPMGPGMGVPGQPMGI PGPBPBEGRYKQKQFQSVFTV 124

QY 109 KSESRVPPSPDAPLPFDVVLVNEQGHYAVTGKPCQVPGVYFAVHATVYRASIQPDL 168

DB 125 TR-QTHQPPAPNSLIRFNAVLTNPQGDYDTSTGKFTCKVPLGYFYTHAS-HTANLCVLL 182

QY 169 VNGGESIAEFQFPGMPPKPSI-SSGAWVRLPEPDQVWVGVGVDPY-IGIYASIKTD 224

DB 183 YKSGVAVT-FCGHTSKTNQVNSGCVLLRLQVSEVW-LAVNDYIMVGIQSS---D 234

QY 225 STFGSFLVYSD 235

DB 235 SVFSGFLFPD 245

RESULT 12

CA18 RABIT STANDARD; PRT; 744 AA.

AC P14282;

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 01-NOV-1990 (Rel. 13, Last annotation update)

DE 01-NOV-1995 (Rel. 32, Last annotation update)

GN COL8A1.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RX MEDLINE=89380199; PubMed=2476437;

RA Yamauchi N., Benya P.D., van der Rest M., Nimmo-Ya Y.;

RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs demonstrate that type VIII collagen is a short chain collagen and contains triple-helical and carboxyl-terminal non-triple-helical domains similar to those of type X collagen."

RT J. Biol. Chem. 264:16022-16029(1989).

CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCENDING MEMBRANE (BASEMENT MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.

CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION WITH ALPHA 2(VIII) TYPE COLLAGENS.

CC -1- PFM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.

CC -1- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLETS ARE PRESENT AT THE C-TERMINUS OF THE TRIPLE-HELICAL REGION. THESE MAY PROVIDE HIGH THERMAL STABILITY OF THIS REGION OF THIS REGION.

CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.

CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; J05042; AAJ1204.1; .

DR PIR; A34246; A34246.

DR InterPro; IPR001073; C1Q.

DR InterPro; IPR000087; Collagen.

DR Pfam; PF00386; C1q; 1.

DR Pfam; PF00391; Collagen; 8.

DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.

DR PROSITE; PS0113; C1Q; 1.

KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation; Glycoprotein; Cell adhesion; Collagen; Signal.

FT SIGNAL 1 20

FT CHAIN 21 744 COLLAGEN ALPHA 1(VIII) CHAIN.

FT DOMAIN 29 117 NONHELICAL REGION (NC2).

FT DOMAIN 118 571 TRIPLE-HELICAL REGION.

FT DOMAIN 572 744 NONHELICAL REGION (NC1).

FT DOMAIN 609 744 C1Q.

SO SEQUENCE 744 AA; 73358 MW; 2A8CE1FE8274E99 CRC64;

Query Match 28.2%; Score 374; DB 1; Length 744;

Best Local Similarity 34.4%; Pred. No. 4.7e-19;

Matches 93; Conservative 36; Mismatches 79; Indels 62; Gaps 8;

QY 24 IPSLC-PGHPGLPCTPGHSSQQLPGRDGDDGAPAGGEGGGRPLP----- 74

DB 475 VPLGLPGKGPGLIPDQGLQGPPIGITYGSGPIGPGKGPGLPDPGPVGVK 534

QY 75 -----GPRGP---GPRGE---AGPAGPMPGAGECVPPS----- 104

DB 535 PGVAGLHGPGRGALGPGQGPGLPGPPGPGRPPGPMPPTPAPQGEYLPDMGLGIDG 594

QY 105 -----AFSAKRESRVP-----PSDAPLPDRVLVNEQGHYAVTGKFT 144

DB 595 VTEPHAYAKKKGKNGPVEWPAFTALTPAPPVGAITKFDRLLYNGRQYVNPQTGIFT 654

QY 145 CQVPGVYFAVAVATYTRASLQFDLVKNGESIA-SFFQFGGMPKPSISGAMVRLPEED 203

DB 655 CEVPGVYFAVAVHCKGKGNVVALFKNNBPVMTYDEVKKGFLDQA--SGSAVLLLRPGD 712

QY 204 QVWVQGVGDYIGIYASIKTDSFSGFLVY 233

DB 713 RVFLQMPSEDAAGLYAGQYVHSSFSFGYLLY 742

RESULT 13

COLL_LEPMA STANDARD; PRT; 419 AA.

ID COLL_LEPMA

AC P98085; Q91080;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Inner ear-specific collagen precursor (Saccular collagen).

OS Lepomis macrochirus (Bluegill).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae; Centrarchidae; Lepomis.

OX NCBI_TaxID=13106;

RN [1]

RX MEDLINE=95167486; PubMed=7863331;

RA Davis J.G., Oberholzer J.C., Burns F.R., Greene M.I.;

RT "Molecular cloning and characterization of an inner ear-specific structural protein."

RT Science 267:1031-1034(1995).

RN [2]

RN CONCEPTUAL TRANSLATION.

RA Gibson T.;

RT Submitted (MAR-1995) to the SWISS-PROT data bank.

CC -1- FUNCTION: FORMS A MICROSTRUCTURAL MATRIX WITHIN THE OTOLITHIC MEMBRANE (PROBABLY).

CC -1- TISSUE SPECIFICITY: SPECIALIZED SECRETORY SUPPORTING CELLS AT THE OUTER PERIMETER OF THE SACCULAR EPITHELIUM.

CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.

CC -----

CC INTRODUCED FOR POSITIONS 391-419 SO AS TO MAXIMIZE THE SIMILARITY WITH OTHER SHORT-CHAIN COLLAGENS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

```

CC      EMBL; U17431; AAAG9978.1; ALT_FRAME.
DR      InterPro; IPR001073; C1q.
DR      InterPro; IPR000087; Collagen.
DR      Pfam; PF00386; C1q; 1.
DR      Pfam; PF01391; Collagen; 3.
DR      PRINTS; PR00007; COMPLEMENTC1Q.
DR      SMART; SM00110; C1Q; 1.
DR      PROSITE; PS01113; C1Q; 1.
KW      Extracellular matrix; Repeat; Collagen; Signal.
FT      SIGNAL 1 19
FT      CHAIN 20 419
FT      DOMAIN 20 57
FT      DOMAIN 58 274
FT      DOMAIN 275 419
FT      DOMAIN 272 419
FT      CARBOHYD 37 37
FT      CARBOHYD 320 320
SQ      SEQUENCE 419 AA; 43634 MW; 570CDB9675FC0F39 CRC64;

Query Match
Best Local Similarity 28.2%; Score 373; DB 1; Length 419;
Matches 96; Conservative 23; Mismatches 99; Indels 38; Gaps 6;

OY      15 GSPPLDNNK-IPLSLCPGHPGLPTPPHHGSGGLPGR-DGRDGDGAPGAPGEGGGR 70
DB      157 GSPGLNGTNGSIIGREGWMPGLAGTGLKGEQGLAGECLQGRGGRGPPGLGEMGLNGT 216
OY      71 PGLPGRGDRGP-----RGEAGPGPTPGAGECV 100
DB      217 DGVKGERGEPGLGKGTGANGPPGPGPGGMAIGKRGKLGKVGPPGPPGPGGSEV 276
OY      101 PPSAFAKRSSESRVPPSPDAPLPEDRVLVNQGSHYDAVTGKTCQVPGVYFAVAHATY 160
DB      277 QIRSNASVGLPPSRSPPLSLPVKFDKLVYNGEHWDPITANKFNVTYFPGVYFSYITVR 336
OY      161 RASLQFDLYKNG-EKSLASFQPPFGMPKASISGGMVLEPEPDVQVGVGDIYGIYA 219
DB      337 NRPVRAALVYNGVRKLRTRDSLYGQDIDA--SNLLALHLVGDQVWLET-LDMNGKYS 393
OY      220 SIKTSTFGSLVYSD 235
DB      394 SSEDSTFGSLIYPD 409

RESULT 14
CAL8_HUMAN STANDARD; PRT; 744 AA.
ID      P27658; Q96D07;
AC      P27658; Q96D07;
DT      01-AUG-1992 (Rel. 23, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Collagen alpha 1 (VIII) chain precursor (Endothelial collagen).
GN      COL8A1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91231001; PubMed=2029894;
RA      Muragaki Y., Mattei M.-G., Yamaguchi N., Olsen B.R., Ninomiya Y.;
RT      "The complete primary structure of the human alpha 1 (VIII) chain and
RT      assignment of its gene (COL8A1) to chromosome 3."
RL      Eur. J. Biochem. 197;615-622(1991).
RP      [2]
RP      SEQUENCE FROM N.A.
PC      Tissue=lung;
RA      Strausberg R.;
RA      Submitted (Sep-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: MAJOR COMPONENT OF THE DESCENDING MEMBRANE (BASEMENT
CC      MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC      -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION

```

```

CC      WITH ALPHA 2 (VIII) TYPE COLLAGENS.
CC      -1- PTH: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC      UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC      -1- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLETS ARE PRESENT
CC      AT THE C-TERMINUS OF THE TRIPEPTIDE-HELICAL REGION. THESE MAY PROVIDE
CC      THE HIGH THERMAL STABILITY OF THIS REGION.
CC      -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC      -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X57527; CAA40748.1; --
CC      EMBL; BC013581; AAH13581.1; --
CC      PIR; S15435; S15435.
CC      DR      Genew; HGNC:2215; COL8A1.
CC      DR      MIM; 120251; --
CC      DR      InterPro; IPR001073; C1q.
CC      DR      InterPro; IPR000087; Collagen.
CC      DR      Pfam; PF00386; C1q; 1.
CC      DR      PRINTS; PR00007; COMPLEMENTC1Q.
CC      DR      PRODOM; PD000007; Collagen; 1.
CC      DR      SMART; SM00110; C1Q; 1.
CC      DR      PROSITE; PS01113; C1Q; 1.
CC      KW      Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
CC      Glycoprotein; Cell adhesion; Collagen; Signal.
FT      SIGNAL 1 28
FT      CHAIN 29 744
FT      DOMAIN 29 117
FT      DOMAIN 118 571
FT      DOMAIN 572 744
FT      DOMAIN 609 744
FT      DOMAIN 262 744
FT      CONFLICT 297 297
FT      CONFLICT 344 344
FT      CONFLICT 382 382
FT      CONFLICT 388 388
FT      CONFLICT 454 454
FT      CONFLICT 464 464
FT      CONFLICT 601 601
FT      CONFLICT 631 631
SQ      SEQUENCE 744 AA; 73364 MW; 2BC1B0955DE2C9A3 CRC64;

Query Match
Best Local Similarity 27.8%; Score 368; DB 1; Length 744;
Matches 90; Conservative 34; Mismatches 84; Indels 62; Gaps 6;

OY      24 IPSLC-DGHPGLPCTPPHHSGGLPGRDGDGAPGAPGEGGGRPLP----- 74
DB      475 VFGILGPGKGBEPGIPEDQGLQGPPIGIGSGSGIGPPGILGPPKGPPLPPLP 534
OY      75 -----AGBCS 99
DB      535 PGVAGLHGPCKPGLALGFGQGGPGLRGPGRGPPGPPVAMPPTPPPGGEVLPDMLGIDG 594
OY      100 VPPSAFAKRSSESRV-----PPSDAPLPEDRVLVNQGSHYDAVTGKFT 144
DB      595 VKPPIAYGAKKGGKNGGPAVEPAFTABLTAPPGAVKKNKLYNGRQYNNPQTGLFT 654
OY      145 CQVPGVYFAVAHATYRASLSQFDLYKNGESA-SFQPPFGGMPKASISGGMVLEBED 203
DB      655 CEVPGVYFAVAHCKGKGVWVALPKNBPWYITTDYKKGFLDQA--SGSAVLLLRGD 712
OY      204 QVWVQGVGDYIGIYASIKTDSFGSLVY 233
DB      713 RVFLQMPSEQAAGLVAGQYVHSSFGSYLLY 742

```


RESULT 15
CA18_MOUSE STANDARD; PRT; 743 AA.
AC 000750; 09D2V4.32, Created)
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(VIII) chain precursor.
CN COL8A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C;
RX MEDLINE=92362626; PubMed=1499564;
RA Muragaki Y., Shiota C., Inoue M., Ooshima A., Olsen B.R.,
RA Nishimura Y.,
RT "Alpha 1(VIII)-collagen gene transcripts encode a short-chain
RT collagen polypeptide and are expressed by various epithelial,
RT endothelial and mesenchymal cells in newborn mouse tissues.";
RL Eur. J. Biochem. 207:895-902(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Atakawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staudt R., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boileau D., Boujuna N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Yashnab-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCENDING MEMBRANE (BASEMENT
CC MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN CALVARIUM, EYE & SKIN OF
CC NEWBORN MICE. ALSO IN VARIOUS EPITHELIAL, ENDOTHELIAL AND
CC MESENCHYMAL CELLS.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66976; CAA47387.1; -;
DR EMBL: X66977; CAA47387.1; JOINED.
DR EMBL: AK018742; BAB31383.1; -;
DR PIR: S23779; S23779.
DR MGI: 88463; Col8a1.

DR InterPro: IPR001073; C1Q.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; C1q; 1.
DR Pfam: PF01391; Collagen; 7.
DR PRINTS: PR00007; COMPLEMENTC1Q.
DR SMART: SM00110; C1Q; 1.
DR PROSITE: PS01113; C1Q; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Collagen; Signal.
FT SIGNAL 1 28
FT CHAIN 29 743
FT DOMAIN 29 117
FT DOMAIN 118 571
FT DOMAIN 572 743
FT DOMAIN 608 743
FT DOMAIN 6 6
FT CONFLICT 85 85
FT CONFLICT 109 109
FT CONFLICT 248 248
FT CONFLICT 313 313
FT CONFLICT 323 323
FT CONFLICT 351 351
FT CONFLICT 596 596
FT CONFLICT 717 719
SQ SEQUENCE 743 AA; 73454 MW; F584D85BD53897F4 CRC64;
Query Match 27.7%; Score 366.5; DB 1; Length 743;
Best Local Similarity 32.8%; Pred. No. 1.5e-18;
Matches 90; Conservative 36; Mismatches 77; Indels 71; Gaps 7;
QY 29 PGHPGLPTGPHGSGQLPGRDGRDGPAPGPKGSGRPLGPRGDP----- 80
DB 470 PGLPGVPLGLGRKGBRGIRGDDGLGPPGIPGIVGSPGIPGPGIRGPKGBRPLGPPPGF 529
QY 81 -----GPRG--AGRAPGPGKSGSVPR----- 103
DB 530 PGVKGKGVAGLHGPPEKPGALPGQGPGLGPPGPPGPPAVVTPSPQGEYLPDMGL 589
QY 104 -----SAFSKR-----SESVP--PPSDAPLPFPRVAVNEGHYDAVT 140
DB 590 GIDGVTPPHVAGKKGKRGSGPAYENPAPFPAELTVPPPPYGPAPYKFKLLYNGRQNNPQT 649
QY 141 GKFTCVPGVYVPAVATVYRASLQFDLVNGESIA-SFFQFGGMPKPKASLGGAMRL 199
DB 650 GITCEVPGVYVPAVAVHCKGAVVVALFENNEPPMYTVDYKKGFLDQ--SGSAVTL 707
QY 200 EPEQYVWVGVGVGYTGYTASITDSTFSGFLY 233
DB 708 RPDQVFLQMPSEDAAGLVAGQYVHSSFSGYLLY 741

Search completed: June 20, 2003, 11:25:04
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 08:45:39 ; Search time 60 Seconds
(without alignments)
834.491 Million cell updates/sec

Title: US-09-944-403-42

Perfect score: 1325
Sequence: 1 MRPLVLLILGLAAGSPPLD.....DSTSGFLVTSMDHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rv1rus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	94.6	243	11 O8R002	O8R002 mus musculi
2	425.5	32.1	240	6 O95M04	O95M04 bos taurus
3	422	31.8	243	6 O95UD7	O95UD7 macaca mla
4	415.5	31.4	675	6 O9N178	O9N178 sus scrofa
5	402	30.3	295	11 O921K4	O921K4 rattus norv
6	395	29.8	294	11 O9D8U4	O9D8U4 mus musculi
7	373.5	28.2	194	6 O95J95	O95J95 canis famli
8	370.5	28.0	744	11 O9D2V4	O9D2V4 mus musculi
9	370.5	28.0	744	11 O921S8	O921S8 mus musculi
10	368	27.8	744	4 O96D07	O96D07 homo sapien
11	363.5	27.4	705	4 O8TEJ5	O8TEJ5 homo sapien
12	314	23.7	245	11 O9DCM6	O9DCM6 mus musculi
13	270.5	20.4	146	11 O9ES30	O9ES30 mus musculi
14	266.5	20.1	196	11 O920N0	O920N0 tamias sibi
15	252.5	19.1	246	13 O91807	O91807 carassius a
16	247.5	18.7	256	13 O91909	O91909 cyprinus ca

17	241.5	18.2	347	4 O961H6	O961H6 homo sapien
18	238.5	18.0	583	4 O96G58	O96G58 homo sapien
19	238.5	18.0	992	4 O9UG76	O9UG76 homo sapien
20	238.5	18.0	1016	4 O9Y6C2	O9Y6C2 homo sapien
21	237	17.9	890	5 O77087	O77087 alvinella p
22	235	17.7	1017	11 O99K41	O99K41 mus musculi
23	234.5	17.7	173	6 O62789	O62789 sus scrofa
24	225	17.0	325	5 O17036	O17036 caenorhabdi
25	223	16.8	281	11 O90XP7	O90XP7 mus musculi
26	222.5	16.8	322	5 O01945	O01945 meloidogyne
27	222	16.8	251	13 O91908	O91908 brachydario
28	222	16.8	1378	5 O97405	O97405 halictus di
29	221.5	16.7	120	6 O77782	O77782 oryctolagus
30	221.5	16.7	248	6 O9T006	O9T006 ovis aries
31	220	16.6	5	O20142	O20142 caenorhabdi
32	219.5	16.6	205	11 O9D0W2	O9D0W2 mus musculi
33	219.5	16.6	326	11 O8R066	O8R066 mus musculi
34	219.5	16.6	381	5 O94399	O94399 caenorhabdi
35	219.5	16.6	589	11 O991L6	O991L6 mus musculi
36	219.5	16.6	1453	11 O63079	O63079 rattus norv
37	218	16.5	182	11 O8R1P2	O8R1P2 mus musculi
38	217	16.4	319	5 O17038	O17038 caenorhabdi
39	217	16.4	1160	4 O14046	O14046 homo sapien
40	217	16.4	1344	3 O93419	O93419 gallus gall
41	217	16.4	1418	6 O28396	O28396 equus cabal
42	217	16.4	1442	11 O62031	O62031 mus musculi
43	217	16.4	1442	11 O62033	O62033 mus musculi
44	217	16.4	1450	13 O911B4	O911B4 cyrtops pyrr
45	217	16.4	1459	11 O62032	O62032 mus musculi

ALIGNMENTS

RESULT 1

ID O8R002 PRELIMINARY; PRT; 243 AA.

AC O8R002; 01-JUN-2002 (TREMBLER, 21, Created)
 DT 01-JUN-2002 (TREMBLER, 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLER, 21, Last annotation update)
 DE Similar to DKFZP586B0621 protein (Hypothetical 25.4 kDa protein).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC023068; AAH23068.1;
 DR EMBL; BC025174; AAH25174.1;
 KW Hypothetical protein.
 SQ SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64;

Query Match 94.6%; Score 1254; DB 11; Length 243;
 Best Local Similarity 93.4%; Pred. No. 5.1e-100;
 Matches 227; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGIPGPHHGSQGLPRDRDRDAPG 60
 1 MRPLVLLILGLAAGSPPLDNDKIPSLCPGHPGIPGPHHGSQGLPRDRDRDAPG 60
 DB 61 ARGERGGRGRLPGRDPPGRGAGAGTGAGGCVPRSAFSAKRSRVPSPD 120
 61 ARGERGGRGRLPGRDPPGRGAGAGTGAGGCVPRSAFSAKRSRVPSPD 120
 QY 121 APLPDRVLVNEQGHYDAVTGKFTCOVGVVYFAVHATYRASLOFDLVKNGESTIASFFQ 180
 121 APLPDRVLVNEQGHYDAVTGKFTCOVGVVYFAVHATYRASLOFDLVKNGESTIASFFQ 180

```

Db      121 TPLPDRVLINEGHFDPTTGKFTQVPGVYFAVHATYRASLOPDLVKNQSIASFQ 180
Qy      181 FFGWMPKPSLISGAMVRLPEPDQVWVGVGVGIVGIVASIKTDSIFSGFLVSDMHS 240
Db      181 YFGWMPKPSLISGAMVRLPEPDQVWVGVGVGIVGIVASIKTDSIFSGFLVSDMHS 240
Qy      241 VFA 243
Db      241 VFA 243

RESULT 2
Q95M04 PRELIMINARY; PRT; 240 AA.
ID Q95M04 PRELIMINARY; PRT; 240 AA.
AC Q95M04 PRELIMINARY; PRT; 240 AA.
DT 01-DEC-2001 (TREMUREL. 19, Created)
DT 01-DEC-2001 (TREMUREL. 19, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE Adipose tissue-specific protein adipo Q.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21369933; PubMed=11382781;
RA Sato C., Yasukawa Z., Honda N., Matsuda T., Kitaajima K.;
RT "Identification and Adipocyte Differentiation-dependent Expression of
RT the Unique Distal Acid Residue in an Adipose Tissue-specific
RT Glycoprotein, Adipo Q."
RL J. Biol. Chem. 276:28849-28856(2001).
DR EMBL; AF269230; AAK58902.1; -.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRODOM; PD000007; Collagen; 1.
DR PROSITE; PS01113; Clq; UNKNOWN 1.
SQ SEQUENCE 240 AA; 26091 MW; C6253BA803B9A668 CRC64;

Query Match 32.1%; Score 425.5; DB 6; Length 240;
Best Local Similarity 41.5%; Pred. No. 7.1e-29;
Matches 100; Conservative 35; Mismatches 81; Indels 25; Gaps 10;

Qy      6 VLLILGLAA-GSPPLDKNIP-SLCPG-HPGI-PGTHGHSQGLPGSDGRDGDARFAP 62
Db      7 LLLILALPSHGDNMEDPPLPKGACAGMAGIPGPHNGT---PGRDGRD-----GTP 57
Qy      63 GEKGEGRPGLPGRPD-----PGRGEAGRPPTGPAAGCSVPSPRSPASAKRS 113
Db      58 GEKGEGRGAGLPGPGEKGTGVMTGAEPRGPGTPGKGEGRGAEAYVRSASFV-GL 116
Qy      114 RVPSPDAPLPDRVLINEGHFDPTTGKFTQVPGVYFAVHATYRASLOPDLVKNQ 173
Db      117 RVTVP-NVPIRFTKIFYNQNHVDSTGTFYCNIGLIVFYSHTIVWKDVKSLEKDK 175
Qy      174 SIASFQFPGWMPKPSLISGAMVRLPEPDQVWVGVGVGIVGIVASIKTDSIFSGFLV 232
Db      176 AVLFYTDYQF-KNVDSAGSVLHLLEVGDVWLQVYGEENNGVYADNVDSITFTGL 234
Qy      233 Y 233
Db      233 Y 233

RESULT 3
Q95J07 PRELIMINARY; PRT; 243 AA.
ID Q95J07 PRELIMINARY; PRT; 243 AA.
AC Q95J07 PRELIMINARY; PRT; 243 AA.
DT 01-DEC-2001 (TREMUREL. 19, Created)
DT 01-DEC-2001 (TREMUREL. 19, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)

```

```

DR      Adiponectin.
GN APML.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OC NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=ADIPOSE TISSUE;
RX MEDLINE=2123234; PubMed=11334417;
RA Hotta K., Funahashi T., Bodkin N.L., Ortmeier H.K., Ariga Y.,
RA Hansen B.C., Matsuzawa Y.;
RT "Circulating concentrations of the adipocyte protein adiponectin are
RT decreased in parallel with reduced insulin sensitivity during the
RT progression to type 2 diabetes in rhesus monkeys."
RL Diabetes 50:1126-1133(2001).
DR EMBL; AF404407; AAK92202.1; -.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRODOM; PD000007; Collagen; 1.
DR PROSITE; PS01113; Clq; UNKNOWN 1.
SQ SEQUENCE 243 AA; 26264 MW; 49A45DAF2B4613FD CRC64;

Query Match 31.8%; Score 422; DB 6; Length 243;
Best Local Similarity 42.3%; Pred. No. 1.4e-28;
Matches 104; Conservative 28; Mismatches 88; Indels 26; Gaps 8;

Qy      6 VLLILGLAAGSPPLDKNIPSLCPGHPG-----LPGTGHSQGLPGSDGRDGDARFAP 60
Db      6 VLLILALPSHGDDTTQGRGVLLPLPKGACTGMMAGIPGPHNGVPPGDRD-----G 59
Qy      61 AFGEKGEGRPGLPGRPD-----PGRGEAGRPPTGPAAGCSVPSPRSPASAKRS 111
Db      60 TPEKGEKEDPGLIPKPGDTGTVGAEPRGPGTQGRKGEPRGSAVYRSASFV-GL 118
Qy      112 ESRVPPSDAPLPDRVLINEGHFDPTTGKFTQVPGVYFAVHATYRASLOPDLVKN 171
Db      119 ETVTVF-NMPIRFTKIFYNQNHVDSTGTFYCNIGLIVFYSHTIVWKDVKSLEK 177
Qy      172 GSS-IASFQFPGWMPKPSLISGAMVRLPEPDQVWVGVGVGIVGIVASIKTDSIFSG 229
Db      178 DRALMTDYQDENNVDA--SGSVLHLLEVGDVWLQVYGEGERNGLYADNDNDSTFTG 235
Qy      230 FLVYSD 235
Db      236 FLVYHD 241

RESULT 4
Q9N178 PRELIMINARY; PRT; 675 AA.
ID Q9N178 PRELIMINARY; PRT; 675 AA.
AC Q9N178 PRELIMINARY; PRT; 675 AA.
DT 01-OCT-2000 (TREMUREL. 15, Created)
DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE Type X collagen.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21015405; PubMed=11130976;
RA Nielsen V.H., Bendixen C., Arnbjerg J., Sorensen C.M., Jensen H.B.,
RA Shukit N.M., Thomsen B.;
RT "Abnormal growth plate function in pigs carrying a dominant mutation
RT in type X collagen."
RL Mamm. Genome 11:1087-1092(2000).
DR EMBL; AF222861; AAF37271.1; -.

```

DR InterPro; IPR001073; Clq.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; Clq; 1.
 DR Pfam; PF01391; Collagen; 8.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR ProDom; PD000007; Collagen; 2.
 DR SMART; SM00110; Clq; 1.
 DR PROSITE; PS01113; Clq; 1.
 DR Collagen.
 KW SEQUENCE 675 AA; 65447 MW; 26397B10310383F9 CRC64;

Query Match 31.4%; Score 415.5; DB 6; Length 675;
 Best Local Similarity 36.1%; Pred. No. 1.7e-27;
 Matches 99; Conservative 34; Mismatches 82; Indels 59; Gaps 7;

QY 15 GSPPLDDNKTPLSLCGHGLPCTPGHHSQGLPGRDGRDGAAPAGEKRGGRPLP 74
 DB 402 GNPGLPGKRGDPRGIGPPGLPQVPGPAKGVPGHNGEAPRGAPGIPGTRGPIGPGLP 461
 QY 75 -----GPRGDPGPRGEA-----GPAFTGPGRCSS 99
 DB 462 GPPGSKGDPGNPGPPGAGIATKGLNGPFGPPGPKGAGBPGLPGRPPGPPGQ-A 520
 QY 100 VPP-----RSASAKRSBSRP-----PPSDAPLPFDRVLVNEQHYAV 139
 DB 521 VPPBGFVKEGQRAFYVSAQGVGMVSAFTVILSKAYPAIGAPLPFDKILNNGOHPK 580
 QY 140 TGKFTQVGVVYFAVHATVYRASLQFDLVKNGESIA-SFPQFGWPKPALSGLGANVR 198
 DB 581 TGIPTCRIPGITYFSHIVKTHAMVGLYKNGTFVMTYDEYVGYLDA--SGSALTD 638
 QY 199 IEPEDQVWVGVGVYIGIYASIKTDSFGFLV 232
 DB 639 LTNDQVWLQLPNAGSNGLYSSSEYVHSSFGFLV 672

RESULT 5
 Q921K4 PRELIMINARY; PRT; 295 AA.
 ID 0921K4;
 AC 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 21, Last annotation update)
 DE Collagen alpha 1 type X (fragment).
 GN COL10A1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=20310874; Pubmed=10853827;
 RA Marks S.A., Lundmark C., Christerson C., Murtz T., Odgren P.R.,
 RA Seifert M.F., Mackay C.A., Mason-Savas A., Popoff S.B.;
 RA "Endochondral bone formation in toothless (osteopetrotic) rats:
 RT failures of chondrocyte patterning and type X collagen expression.";
 RL Int. J. Dev. Biol. 44:309-316 (2000).
 DR EMBL; AJ31848; CA10518.1; -
 DR InterPro; IPR001073; Clq.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; Clq; 1.
 DR Pfam; PF01391; Collagen; 3.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR SMART; SM00110; Clq; 1.
 DR PROSITE; PS01113; Clq; 1.
 FT NON_TER 1
 SQ SEQUENCE 295 AA; 30012 MW; FF43B1548028813E CRC64;

Query Match 30.3%; Score 402; DB 11; Length 295;
 Best Local Similarity 36.7%; Pred. No. 9.3e-27;
 Matches 94; Conservative 31; Mismatches 79; Indels 52; Gaps 7;

QY 25 PSLCPGH-----PGLPTGHHGSGQLPGRDGRDGAAPAGE-----KCE 67
 DB 41 PKGVPGHNGEAPRGGRGPIGTRGPIGPGRGSKDPGKPPGAPGAGIVTKGLNRP 100
 QY 68 GGRPGLGPR---GDPGREGAGPAGPTGPGRCSP----- 101
 DB 101 AGPPRGPRGRGHTGEBGLPGPPGPBPBPSSQAVIPDFTYSGORPLSGMPLYSANOG 160
 QY 102 ----PSAPAKRSBSRVPSPDAPLPFDRVLVNEQHYDAVTKFTQVGVVYFAVHA 157
 DB 161 VTGMPVSAFTVILSKAV--PAVGAIPFDEILNROOHDRSGITCTKIPGITYSYH 218
 QY 158 TVYRASLQFDLVNNG-ESIASFPQFGWPKPALSGLGANVRLEPEDQVWVGVGVYIG 216
 DB 219 HVGHTHVGVLKNGGPTVMTYDEYSKGYLDA--SGSALMELTENDQVWLQLPNBSNG 276
 QY 217 TVASIKTDSFGFLV 232
 DB 277 LYSSEYVHSSFGFLV 292

RESULT 6
 Q9DBU4 PRELIMINARY; PRT; 294 AA.
 ID 09DBU4;
 AC 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE 1810033K05RIK.
 GN 1810033K05RIK protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okasaki Y., Gojopori T., Bono H., Kanukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Stanbii F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G.,
 RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Hynes-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 DR EMBL; AK007683; BAB25187.1; -
 DR WGI; MGI:1916433; 1810033K05RIK.
 DR InterPro; IPR001073; Clq.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; Clq; 1.
 DR Pfam; PF01391; Collagen; 2.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR SMART; SM00110; Clq; 1.
 DR PROSITE; PS01113; Clq; 1.
 FT NON_TER 1
 SQ SEQUENCE 294 AA; 30865 MW; 6D3905AE7C19B6FA CRC64;

Query Match 29.8%; Score 395; DB 11; Length 294;
 Best Local Similarity 35.9%; Pred. No. 3.7e-26;
 Matches 94; Conservative 31; Mismatches 93; Indels 44; Gaps 6;

```

QY 3 PLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGPHHSGGLPGRDGRDGRGAP 62
DB 41 FOLVCSLPG-----PGPPGPPGAPGSSGVVGRGPGKGGDDDDRDSD 87
QY 63 GEGKGGGPR---GLPGRGDPGRGAGAPG---TGPAGCSVPRAFAKRSSESNVP 116
DB 88 GEGGPGGTGNRGKGGPKAGATGAGGPRGKGVSGTPGKGTGCKGPKKKEPRLP 147
QY 117 PPSDA-----PLPFDRLVNEQGHYDAVTGKTCQVPGVYFA 154
DB 148 GPCGCGSSRAKSAFSAVAVTKVSRERLPIKFDKILINBEGHYNASSGKFCVSPGIYFT 207
QY 155 VHAATYRASLOPDLVKNESIASFPQFGGPKRPSLSGAMVRLPEPQVAVVGVGDY 214
DB 208 YDITLANKHLAIGLVNNGQYRIKTPDANTGNHVA--SGSTILAKGDEVLQIFYSBQ 265
QY 215 IGI-YASIKTSTFSGELVYSD 235
DB 266 NGLFDPYWTDSLFTGFLIYAD 287

```

RESULT 7

```

ID 095J95 PRELIMINARY; PRT; 194 AA.
AC 095J95;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Adiponectin (Fragment).
GN APM1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxId=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADIPOSE TISSUE;
RA Kabir M., Ananthanarayan S., Ionut V., Kim S.P., Van Citters G.W.,
RA Dea M.K., Bergman R.N.;
RT "Regulation of Adiponectin gene expression in the fat-fed dog.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF117206; AAL09702.1; -
DR InterPro: IPR001073; Clq.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; Clq; 1.
DR Pfam: PF01391; Collagen; 1.
DR ProDom: PD000007; Collagen; 1.
DR PROSITE: PS01113; Clq; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 194
SQ SEQUENCE 194 AA; 20890 MW; 3AA3D947D187AFA CRC64;

```

Query Match 28.2%; Score 373.5; DB 6; Length 194; Best Local Similarity 42.3%; Pred. No. 1.6e-24; Matches 90; Conservative 22; Mismatches 66; Indels 35; Gaps 9;

```

QY 8 LLLGLAAGSPPLDNNKIPSLCPG-HRGLPTGPHHSGGLPGRDGRDGRGAPAPGEGK 66
DB 4 VLLPLPFGA-----CPGMAAGIPGHPGNGT---PGRGRD-----GTPEKG 43
QY 67 EGRPGPLGPRGD-----PGRGAGPAGPTGPAAGCSVPRAFAKRSSESNVP 117
DB 44 EKGPGGLVGRKGDGTGVTGVEGRGRGPRGPRGKGRGEGSAVYHRSASFV-GLSRTIV 102
QY 118 PSDAPLPFDRLVNEQGHYDAVTGKTCQVPGVYFAVHAIVYRASLOPDLVKNESIAS 177
DB 103 P-NVPIRPTKIFYNLQNHYDGTGKFCNIPGLYFSTHITVYLKDVASLYKDK--AM 159
QY 178 FFGPRGWPXDA-SLSGAMVRLPEPQVAVVGVGDY 209
DB 160 LFTYDQYQKAVDQASGVLLHLEVGDDVWLYQV 192

```

RESULT 8

```

ID 09D2V4 PRELIMINARY; PRT; 744 AA.
AC 09D2V4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Collagen, type VIII, alpha 1.
GN COL8A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batelov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barin G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Gustincich S., Hill D., Hofmann M., Mazza J., Mombarts P.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK018742; BAB31383.1; -
DR MGI: 88463; Col8a1.
DR InterPro: IPR001073; Clq.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; Clq; 1.
DR Pfam: PF01391; Collagen; 7.
DR PRINTS: PR00007; COMPLEMENTC1Q.
DR SMART: SM00110; Clq; 1.
DR PROSITE: PS01113; Clq; 1.
SQ SEQUENCE 744 AA; 73581 MW; C659BDCBCEDEB9C CRC64;

```

Query Match 28.0%; Score 370.5; DB 11; Length 744; Best Local Similarity 34.6%; Pred. No. 1.4e-23; Matches 93; Conservative 32; Mismatches 83; Indels 61; Gaps 7;

```

QY 24 IPSLC--RGHPLPTGPHHSGGLPGRDGRDGRGAGARGEKGGKGRPLP----- 74
DB 476 VGLLSPKSGEPRIPEDQGLQGPPIGLVGS8GPIGPGRGPRGPRGPRGPRGK 535
QY 75 -----GPRGDPGRGAGAPGPTGP-----AGCSV 100
DB 536 PGVAAGHPRPKRGAALGPGQGGLGPPGPPGPPGPPVAMPPTSPGSEYLPDMGIGIDGV 595
QY 101 PRSAFSAKR-----SESRVP--PRSDAPLPFDRLVNEQGHYDAVTGKTC 145
DB 596 KPPHYVAGKGRGGRGAPAYEMPAFTALVPPPPVGAIVFDTLLNGQNINPQTGITC 655
QY 146 QVGVYVFAVHAIVYRASLOPDLVKNESIAS-SFPQFGGWPXDA-SLSGAMVRLPEPQ 204
DB 656 EIPGVYVFAVHAIVYRASLOPDLVKNESIAS-SFPQFGGWPXDA-SLSGAMVRLPEPQ 713
QY 205 VVAVGVGVYIGIVASIKTSTFSGELVY 233
DB 714 VFLQMPSEDAAGIYAGQYVHSSPSGYLLY 742

```


Db	643	NPATYTYDEYKKGYLDDA--SSGAVLQLRPNDQVWQMPQDANGXSTYTHSSFGCF	700
Qy	231	LV 232	
Db	701	LL 702	
RESULT 12			
Q9DCM6			
ID	Q9DCM6	PRELIMINARY;	PRT; 245 AA.
AC	Q9DCM6:		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	Complement component 1, q subcomponent, alpha polypeptide.		
OS	C10A.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OC	NCBI_TaxID=10090;		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=C57BL/6J; TISSUE=KIDNEY;		
RC	MEDLINE=21085660; PubMed=11217851;		
RX	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi Y., Fukuda S.,		
RA	Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,		
RA	Saito T., Okazaki Y., Gojodori T., Bono H., Kanakawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., Kinsle G., Quackenbush J.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nakai T., Tomita M., Wagner L., Washio T.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Baldarelli R., Barsh G.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Carminci P., de Bonaldo M.F.,		
RA	Blake U., Boffelli D., Bojunga N., Carninci P., Fujita M., Gariboldi M.,		
RA	Brownstein S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombauts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,		
RA	Hayashizaki Y.;		
RT	Functional annotation of a full-length mouse cDNA collection.;		
RL	Nature 409:685-690(2001).		
RP	SEQUENCE FROM N.A.		
RA	Straussberg R.;		
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AK002655; BAB2262.1; -		
DR	EMBL; BC002086; AAH02086.1; -		
DR	MGD; MGI:88223; C1ga.		
DR	InterPro; IPR0101073; C1g.		
DR	InterPro; IPR000087; Collagen.		
DR	Pfam; PF00366; C1g; 1.		
DR	Pfam; PF01391; Collagen; 1.		
DR	PRINTS; PR00007; COMPLEMENTC1D.		
DR	SMART; SM00110; C1Q; 1.		
DR	PROSITE; PS01113; C1Q; 1.		
DR	SEQUENCE 245 AA; 25974 MW; 41C206D49592020 CRC64;		
Qy	Query Match	23.7%; Score 314; DB 11; Length 245;	
Db	Best Local Similarity 35.3%; Pred. No. 2,7e-19;		
Db	Matches 79; Conservative 24; Mismatches 85; Indels 36; Gaps 6;		
Qy	29	PGHGLPPTPTGHHSSGSLPGDDGDDGDAFGAPGKKGEGCGFGLGPGRGDGPGRGAGP	88
Db	36	PGNPRPFRPGFLKRGKGGPPGAAG--RTGIRGRFGDGSGSPGKRGNGVGLPSPSPPLD	93
Qy	89	AGPTPAGECGVP-----PRSAFSAKRSRSRVPPSPDADLPFDRLVINEGQHYDAVTK	142
Db	94	SGPQGLKVKNGKNGIRIDQPPARFSAIRQN---PMITGANVITFDKVLTVQESPYQWHTGR	150
Qy	143	PTCGVPGVYVFAVNAVTVRAALQDPLVANGSSIASPFQFGWPKPA-----	189

Ddb		151 FICAVGEYFF-----NFQIVIKMPLCLFIKSSSGGQRDSISFNNTNKKLFQ	199
Oy		190 SLGGAMVRLREPEDOVMVGVDYGIFASIKTSTFPGFLVY	233
Dd		200 VLAGGTVLQLRGDEVIWKDPAGK-RITYCGEADSIIFSGLIF	242
 RESULT 13 O9ES30 PRELIMINARY; PRT; 246 AA.			
ID O9ES30 AC O9ES30:	PRT;	246 AA.	
DT 01-MAR-2001 (TREMBLrel_16,	Created)		
DT 01-MAR-2001 (TREMBLrel_16,	Last sequence update)		
DT 01-DEC-2001 (TREMBLrel_19,	Last annotation update)		
DR Collagenous repeat-containing	sequence of 26kda protein.		
CN COR8 COR826.	Mus musculus (Mouse).		
OC Eukaryota; Metazoa; Chordata;	Cranialia; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Rodentia;	Sciurognathi; Muridae; Murinae; Mus.		
OX NCBI_TaxID=10090;			
[1] -----			
RP SEQUENCE FROM N.A.			
RA MEDLINE=21264842; PubMed=11071891;			
MA Maeda T., Abe M., Kurisu K., Jikko A., Furukawa S.;			
RT "Molecular cloning and characterization of a novel gene, COR826,			
encoding a putative secretory protein and its potential involvement in			
skeletal development."			
J. Biol. Chem. 276:3628-3634(2001).			
EMBL AF246265; AAC33704.1; -			
GCD MGJ:1932136; Cors.			
DR InterPro: IPR001073; Clq.			
DR Interpro: IPR000087; Collagen.			
PFam PF00386; C1g; 1.			
DR Pfam: PF01391; Collagen. 1.			
DR PRINTS: PRO0007; COMPLEMENTC1Q.			
DR SMART: SM00110; C1Q; 1.			
DR POSTSITE: PS01113; C1Q; 1.			
SQ DR POSITE: 246 AA; 26828 MM; 4ZA4B1B3EPFA87E CRC64;			
Query Match	20.4%; Score 270.5; DB 11; Length 246;		
Best Local Similarity	31.8%; Pred. No. 1.5e-15;		
Matches 78; Conservative 33; Mismatches 107; Indels 27; Gaps 6			
Oy	4 LTVLLTLTGIA-----AGSPDLDNKIPSLCRGHRLDGTRGHHSOGILPPRDR	52	
Dd	11 LLPIALFPCLQGDERYMESPQAGGI.PPCSCKCCHGDYGFGRYGOAPPGPBGPIPAHGN	70	
Oy	53 DERODGAPGARPKRGRCARGLP.GRPGRDGPGRGAGPAPGPCAGECVPRSAFSAKRSE	112	
Dd	71 NGNNNGATGHEHGAKG----KGDKGDLGPPRGEGQHCPKEKGYPVVPVELQIAFMASL	124	
Oy	113 SRVPPPSDA.LPFDRLVNBOGHAYDATGKTTCOVPGYYTFVAV---HAIFYRASLOFDL	168	
Dd	125 ATHTFSNQNSGIIIFSSVEINIGNPFOVTMGREFCAPVSQVFPTFSMHMKBEVEEVYV--L	182	
Oy	169 VKNGECIASFGPOFGMPKPASLSGAAVRLREPEDQMVGWGVDDYIGIYAISKTDSTFS	228	
Dd	183 MHNGNVTFMSYSYETKGS DTS-SNHVAYLAKADEVALRMNGC---ALMGDHQRFSTFA	238	
Oy	229 GFLLVY 233		
Dd	239 GFLLF 243		
 RESULT 14 O9ZON0 PRELIMINARY; PRT; 196 AA.			
ID O9ZON0 AC O9ZON0:	PRT;	196 AA.	
DT 01-DEC-2001 (TREMBLrel_19,	Created)		
DT 01-DEC-2001 (TREMBLrel_19,	Last sequence update)		
DT 01-JUN-2002 (TREMBLrel_21,	Last annotation update)		
HP-20.			

```

DR PROSITE; PS00615; C_TYPE LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 1.
KM Lectin; signal.
FT NON TER 1 1 POTENTIAL.
FT SIGNAL <1 13 S -> F.
FT VARIANT 145 145
SQ SEQUENCE 246 AA; 25709 MW; AB692282D289D0D5 CRC64;

Query Match 19.1%; Score 252.5; DB 13; Length 246;
Best Local Similarity 31.9%; Pred. No. 5.2e-14;
Matches 80; Conservative 33; Mismatches 73; Indels 65; Gaps 14

QY LVLTLILGLAGSPPLDNNKIPLSLCPGHLPGTTPGHHSQGLPGRDGRDGRDARGARG 63
DB 2 LILQFALQLLDGAEHQNNL-----CPAYGVGPTPGHN-----GLPGRDGRDGRDGAIGPKG 53
QY 64 EKGG-----GRPGLPGRPGDPGRGEGAPRPTGPAEGCSVPRSAFSAKRSERVPP 118
DB 54 EKSGSGVSVQGPGRKAGPPTGAGEKGERGSPGSPGSSVLSLSKIYQLKAKI----- 110
QY 119 SDAPLPDRVLVNEQGHYDAVTGKFTCVPGVYFPAVHATVYRASLQFDLVKNGESIAF 178
DB 111 ---ATPEK--VSSVCHFRKV-----GQKVIITDGVVG-----NFD-----QGLKSC 146
QY 179 FQFPGWPKRPASLSGGAMV--RLRPDQVVQV-----GVGD---YIGIYASIKTDSFF-- 227
DB 147 MEF-----GGTWSPTSAMENQALLKLVLVSSGLSKRPYIGV--TDRKTEGQFVD 194
QY .228 --SGFLVYSDW 236
DB 195 TEKGQLTFNNW 205

Search completed: June 20, 2003, 11:24:16
Job time : 63 secs

```